

NSL.7

Linguistic Studies in the Non-Slavic Languages of the
Commonwealth of Independent States and the Baltic
Republics

Edited by
HOWARD I. ARONSON
University of Chicago

Chicago Linguistic Society
The University of Chicago
Chicago, Illinois
1994

NSL.7

Linguistic Studies in the Non-Slavic Languages of the
Commonwealth of Independent States and the Baltic
Republics

Edited by
HOWARD I. ARONSON
University of Chicago

Chicago Linguistic Society
The University of Chicago
Chicago, Illinois
1994

Published 1994 by
THE CHICAGO LINGUISTIC SOCIETY.NSL
The University of Chicago
1130 East 59th Street
Chicago, Illinois 60637

Each contribution copyright ©1994 by its author, All rights reserved.

The editor and the publisher assume no responsibility for statements
of fact or opinion by the authors.

ISBN 0-914203-44-4

Printed in the United States of America

TABLE OF CONTENTS

| | |
|--|-----|
| <i>Aronson, Howard I.</i> "Indirect objects" and datives in Georgian. | 1 |
| <i>Colarusso, John</i> . Microcosmos: the Circassian verb. | 15 |
| <i>Comrie, Bernard</i> . An evaluation of Chukchee orthography. | 55 |
| <i>Dyer, Donald L.</i> Russian and Romanian intertwined: the legacy that is Moldavian. | 65 |
| <i>Friedman, Victor</i> . Ga in Lak and the 3 "there's." | 79 |
| <i>Harlig, Jeffrey</i> . The status and maintenance of the Mordvinian languages. | 95 |
| <i>Harris, Alice C.</i> Ergative-to-nominative shift in agreement: Tabassaran | 113 |
| <i>Jacobs, Neil G.</i> On a structural 'fifth column' in sociolinguistic change: the diffusion of a standard Yiddish feature in Yiddish dialects. | 133 |
| <i>King, J. R. P.</i> Dialect features in Soviet Korean publications from the 1920s. | 151 |
| <i>Lallukka, Seppo</i> . The status of the Volga-Finnic languages in the Soviet general education school. | 185 |
| <i>Manning, H. Paul</i> . Split locativity in Old Georgian and Svan. | 211 |
| <i>Polinsky, Maria S.</i> Relativization in Chukchee. | 241 |
| <i>Pugh, Stefan</i> . The assimilation of Russian verbal lexicon in Mokša Mordvinian. | 263 |
| <i>Rounds, Carol</i> . The role of the zero accusative in Komi folk tales. | 279 |
| <i>Schwink, Frederick W.</i> On the lexicalization of Classical Armenian vowel epenthesis. | 287 |
| <i>Testen, David</i> . A feminine/diminutive suffix in early Ossetian. | 299 |
| <i>Tuite, Kevin</i> . Aorist and pseudo-aorist for Svan atelic verbs. | 319 |
| <i>Verschik, Anna</i> . Code-switching and the Estonians of Narva. | 341 |
| <i>Vovin, Alexander</i> . On accent in Chuvash. | 347 |

PREFACE

Most of the papers in this volume were originally presented in abbreviated and, often, in preliminary form at the Seventh International Conference on the Non-Slavic Languages of the USSR, held at the University of Chicago in May 1991. This was the last of our biennial conferences to be held in the lifetime of the late Soviet Union. Now the languages that we deal with belong to fifteen newly independent republics. Yet the influence that Russian exerted over them will continue to be felt in the near future. And many of these languages are little known and little studied outside of the homelands. It is our hope that these volumes will continue to provide a home for the study of languages that otherwise would not have a home and to provide a venue for studies that deal with the various languages of the area in their interrelations.

This volume is entitled *NSL.7* because it is the seventh volume in the series of papers dealing with the non-Slavic languages of what is now the former Soviet Union. Papers from the first conference appeared in: Paul R. Clyne, William F. Hanks, and Carol L. Hofbauer, editors. 1979. *The elements: a parasession on linguistic units and levels, including papers from the Conference on the Non-Slavic languages of the USSR*. Chicago: Chicago Linguistic Society. Pp. 297-481. Papers from the second conference comprised volume 5 of *Folia Slavica* (1982) and papers from the third conference comprised volume 7, numbers 1 and 2, of *Folia Slavica* (1984). Papers from the fourth conference are forthcoming in a volume to be published by Slavica Publishers of Columbus, Ohio. Papers from the fifth conference appeared in: Howard I. Aronson, ed. 1989. Chicago: Chicago Linguistic Society and papers from the sixth conference appeared in: Howard I. Aronson, editor. 1992. *The non-Slavic languages of the USSR: Linguistic studies. New series*. Chicago: Chicago Linguistic Society.

It is my very pleasant duty to express my gratitude to those who have helped to make this volume possible. First of all, I should like to thank those who served as referees for the papers submitted to this volume. Thanks, as always, are due to Bill J. Darden, conference co-organizer, to Philip Gossett, Dean of the Division of the Humanities of the University of Chicago, for guaranteeing funding for the publication of this volume, and the officers of the Chicago Linguistic Society for publishing this volume. Thanks again go to David Testen,

who has provided invaluable aid to the preparation of this and preceding volumes.

Bill Darden and I are honored to dedicate this volume to Professor František Svejkský on the occasion of his retirement from the Department of Slavic Languages and Literatures of the University of Chicago. Although, despite the extent of his scholarly interests, to the best of our knowledge he has never specifically dealt with the areas covered by this volume, he nevertheless epitomizes the breadth and scholarly rigor that we think characterizes the contributions to this volume.

Howard I. Aronson
January 1994

DATIVES AND INDIRECT OBJECTS IN GEORGIAN

Howard I. Aronson
University of Chicago

I have attempted to show elsewhere (Aronson 1979 and Aronson 1991) that morphological and syntactic categories cannot be defined universally, but need to be defined for each language in which they occur in language-specific terms. Further, such definitions should be, whenever possible, in terms of the *internal* structure of the given language, rather than relying on language-external, real-world concepts.

The problem of determining whether the *syntactic* categories *subject*, *direct object*, and *indirect object* are to be found in modern literary Georgian, and, if they do exist, how to define them is particularly resistant to resolution. If it is true that the concept of grammatical subject is not relevant for literary Georgian (for arguments supporting this view, see Aronson forthcoming and Tuite 1988), it follows logically that the concept of the *direct object* is not relevant either. Similarly, the modern Georgian counterpart of a dative argument that was best viewed as an indirect object in Old Georgian has been identified as a grammatical subject, direct object, and indirect object¹; i.e., there are also serious problems in positing an indirect object as a clear syntactic category in modern Georgian.

One aim of the present paper is to attempt to distinguish between various types of use of the dative case with respect to its marking or lack of marking in the verb form. In this respect, we must differentiate between *associated datives*, i.e., datives that are correlated with a formal marking in the verb form and *unassociated* or *non-agreeing datives*, i.e., those datives which are *not* associated with a correlated formal marking in the verb form. For the sake of simplicity (but at the cost of a good deal of terminological inaccuracy, since these can correspond to both grammatical subjects and indirect objects in other languages), I propose to call the former *indirect objects* and the latter, *non-agreeing datives*. I shall also attempt to explain why transitive (I. conjugation) and activity verbs (III. conjugation) do not allow an indirect object in the perfect series while indirect objects are possible

¹In Old Georgian the verb form never showed number agreement with the dative argument. For an attempt to show that the dative argument in the present series of transitive verbs was probably originally not a direct object, see Aronson 1979.

with intransitive verbs (II. conjugation)

The notion of indirect object: It is necessary to distinguish between concepts found on the paradigmatic level, e.g., 'nominative,' 'accusative,' 'dative,' etc., in languages with case systems and concepts existing on the syntagmatic level, e.g., 'subject,' 'direct object,' 'indirect object,' etc. There must be a structural reason for singling out a given noun phrase as an "indirect object." In a language such as French the fact that only certain noun phrases (certain of those consisting of the preposition *à* plus substantive) can be marked in the verb phrase with the clitics *lui*, *leur* allows us to set up such noun phrases as indirect objects, while those noun phrases which can be replaced by *le*, *la*, *les* in the verb phrase will be direct objects. Similarly, in a language like Bulgarian, those phrases headed by the preposition *na* which can be marked in the verb phrase by the clitics *mi*, *ti*, *mu*, *i*, *im*, *si* are indirect objects, while those noun phrases marked in the verb phrase by *me*, *te*, *go*, *ja*, *gi* are direct objects. In both French and Bulgarian, there are also verb phrase clitics which formally neutralize the distinctive marking of the opposition indirect object/direct object:

| (1.) | FRENCH | | BULGARIAN | |
|------|--------|-------|-----------|-------|
| | id.o. | d.o. | id.o. | d.o. |
| 1. | me | | 1. mi | me |
| 2. | te | | 2. ti | te |
| 3. | lui | le/la | 3. mu/i | go/ja |
| 4. | nous | | 4. ni | |
| 5. | vous | | 5. vi | |
| 6. | leur | les | 6. im | gi |

In yet other languages, there seems to be little necessity of positing a special, separate category of indirect object. So, contemporary Russian grammars distinguish between a direct object and *kosvennye ob"ekty* (lit., 'indirect objects'), but these latter can be in any case other than accusative or be a prepositional phrase; the dative object is not singled out for special treatment.² And it is true that in Russian there is little

²Cf. Filin 1979 s.v. *dopolnenie* (p. 76), where a *prjamoe dopolnenie* (direct object) is stated to be expressed by a substantive in the accusative without preposition, while a *kosvennoe priglagoľ'noe dopolnenie* (indirect ad-verbal object)

difference between an object in the dative and, e.g., an object in the instrumental; for neither is there a special marking in the verb form (as with the indirect objects of French and Bulgarian), nor is there any other special treatment of dative objects.

When we turn to Georgian, it is clear that there is a special treatment of dative "objects," a treatment not found with "objects" in the genitive, instrumental, or adverbial cases or with postpositional phrases. This is the use of special morphemes in the verb form which are correlated with *some* dative arguments in the sentence. There are, it is true, many instances in which the formal opposition in the verb between direct object and indirect object is neutralized, and there are very many instances in which there is no formal distinction possible between the presence and absence of a third person object. Further, there is a significant number of instances where what I shall be calling "indirect object markers" do not, in fact, mark the presence of an indirect object. I shall disregard such situations in what follows.

The existence of "indirect objects" as a syntactic category in Georgian is said to be supported by the following person markers correlated with dative actants, as traditionally given in Georgian grammars:

(2).

| I. | | II. | | III. | | IV. | |
|----------|----------------------|-----|----------|------|----------|-----|----------|
| m- | gv- | mi- | gvi- | ma- | gva- | me- | gve- |
| g- | g-...-t ³ | gi- | gi-...-t | ga- | ga-...-t | ge- | ge-...-t |
| h-/s-/Ø- | | u- | | a- | | e- | |

(Markers of type IV are extremely rare with I. and III. conjugation verbs. They are found in the aorist series screeves of the verb *mi/mo-cema* 'give' (e.g., *momeci* 'give it to me,' *mogeci* 'I gave it to you'). In addition, *e-* is found with such verbs as *ejebs*, *ejiebs* 'look for.' Otherwise this marker is generally restricted to dative arguments in II and IV conjugation verbs and to the dative argument in the pluperfect of I. and III. conjugation verbs. In the following discussion we shall

"can be expressed by forms of the indirect cases of the substantive with or without prepositions" (the "indirect" [*kosvennye*] cases are all cases except the nominative).

³In what follows we shall disregard the conditions leading to the loss of the plural marker *-t* when associated with the 2nd person object marker.

ignore type IV.)

***h-* vs. *u-* vs. *a-*.** Given the formal differences in the marking of indirect objects illustrated by (2.) above, the question arises: are there any semantic, functional distinctions correlated with the different markers of indirect objects? Before this question can be answered, however, it must be ascertained that the vowels *-i-*, *-a-* in such sequences as *m-i*, *g-i-*, *m-a-*, *g-a-* are in fact markers of dative relationships. It seems clear that the *-i-* of *m-i*, *g-i-*, *gv-i-* is correlated with a dative object because of these forms' paradigmatic relationship with the 3d person marker *u-*. In the case of *-a-*, it would be hard to view the preradical vowel as a marker of a dative argument. To use Tschenkéli's terminology (GDW 1965:xxvii), the preradical vowel *a-* can be a marker of T¹ (transitive in neutral version), T⁴ (transitive in superessive version), or T⁵ (transitive with 2 objects).⁴ Only T³ and T⁵ presuppose the possibility of the presence of a dative object. But the superessive indicates *either* the presence of a dative object *or* the presence of a postpositional phrase (usually with *-ze*); examples:

(3.) a. Mamam da-**a**-çera misamarti konvert-**s**. 'Father wrote the address on the envelope (DAT).'

b. Mamam da-**a**-çera misamarti konvert-**ze**. 'Father wrote the address on the envelope (postpositional phrase).'

(Tschenkéli 1958:I-395)

Similarly, when the 'neutral version' form of a verb has the preradical vowel *a-*, the vowel of the superessive does not clearly mark the presence of a dative argument:

(4.) a. Me v-**a**-sxam ġvinos. 'I pour the wine (d.o).'

b. Me v-**a**-sxam çqals qvavileb-**s**. 'I pour water (d.o.) on the flowers (id.o).'

(Tschenkéli 1958:I-395)

It is only when the 'neutral version' form does not have the preradical vowel *a-* and the T⁵ form does that the preradical vowel indicates the

⁴For Tschenkéli (GDW), T⁶ is marked either by the preradical vowel *a-* or by the *h*-series indirect object markers.

presence of a dative argument. (Note that this is, again, a paradigmatic, rather than syntagmatic, indication of the presence of the dative argument.) Example:

- (5.) T¹. Samğvdeloebašiac gavrcelda seni da aman jlier dabla *da-s-çia*⁵ samğvdeloeba. 'In the clergy a disease became widespread and this brought the clergy (d.o.) down very low.' (KEGL s.v. *da-s-çevs*).
- (6.) T⁵. Mosamsaxurem ... stxova paṭrons, erti siṭqva *da-m-a-çie* Kuciḡostan, tavisī kali momatxvoso. 'The employee asked [his] boss: "Put in a good word for me (id.o.) at Kuciḡo's, he is to marry his woman to me.'" (KEGL s.v. *daacevs*)

Thus, we must conclude that for I. conjugation verbs in the present/future and aorist series, the markers of an indirect object are only:

(7.)

| | |
|----------|------------|
| m- / gv- | mi- / gvi- |
| g- | gi- |
| h-/s-/Ø- | u- |

The first group is, however, identical to the markers of direct objects (except in the 3rd person), but, since the use of *h-* and *s-* is far from consistent (Aronson 1992:6), we can conclude that, in fact, the first group contains *object markers* and not *indirect object markers*.

We can now turn to the question of whether there is a semantic or functional difference between the *h*-series markers when they mark a dative object and the *u*-series markers of indirect objects. Harris (1981:93ff.) proposes a distinction between *initial indirect objects*, which are "basic and unmarked" and *benefactives* (the traditional "objective version" (*sasxviso kceva*)), which are "derivative and marked" (Harris 1981:93). From her examples there is a clear implication that forms with *u*-series markers are benefactive and forms with the

⁵Despite the fact that there is the O³ marker *s-* here, there is no dative object associated with the verb form here and the form is treated by GDW (s.v. *çev-*) as T¹.

h-series markers and no preradical vowel mark initial indirect objects. Yet there are many examples of verbs which appear to be without an "initial indirect object" which take *h*-series objects as well as examples of forms with *u*-series objects which appear to be just as obligatory as the indirect object associated with a verb such as *mişçera* 'write s.o.'

- (8.) a. *moṭexs* 'break sthg. off' \\\ *mo-s-ṭexs* 'break sthg. off sthg.'
 b. *utxari/utxra* 'said sthg. to s.o.' **txari*/**txra*

But, there are numerous examples where there is a formal contrast between forms of the same verb with *h*-series markers and with *u*-series markers. Very often in such circumstances the form with the *h*-series markers relates a part to a whole, while the form with the *u*-series markers have benefactive meaning. Examples:

- (9.) a. *Babucam ... pūri amoiḡo, ... coṭa Maros mouṭexa*. 'Babuca took out the bread ... he broke off a little for Maro.' (KEGL s.v. *mouṭexs*)
 b. *Berijem ... moṣṭexa cecxlze ɕartul pičxs ɕoṭi*. 'Berije broke off a branch of the brushwood on the fire. (KEGL s.v. *moṭexs*)

However, this verb form also occurs without any indirect object:

- c. *Ori nedli ɕaro moṭexa*. 'He broke off two young ears of corn.' (KEGL s.v. *moṭexs*)

Yet we also find examples where the *u*-series and *h*-series markers both occur, with apparently little semantic difference; such is the case with *ačveneb*s/*učveneb*s:

- (10.) a. *Saati dros ačveneb*s. 'A clock shows the time.'
 b.. *...binas gačveneb*. 'I will show you the apartment.'
 c. *Gačirveba mičvene da gakcevas gičveneb*o. 'Show me

poverty and I will show you fleeing.'

- d. Buxris tavze dadgmuli saati rvas učvenebda.. 'The clock on top of the fireplace showed 8:00.'

In sentences (a.) and (d.) there is no indirect object, while sentences (b.) and (c.) both have indirect objects. There is then, little or no difference between these two verb forms, either syntactically or semantically. From a structuralist point of view, one would be surprised to find such distinct forms not correlated with distinctions in meaning, but the tendency for different forms to have different functions in a language is just that, a tendency, not a law. This, though is clearly an area requiring further investigation.

Unassociated datives. As noted above, positing the existence of the category of indirect object in a given language requires more than the mere existence of dative "objects." (It is on such a basis that the so-called accusatives of time of many Indo-European languages are not regarded as direct objects.) In Georgian, the presence of overt "indirect object markers" in the verb form is sufficient to establish the syntactic category of indirect object. On this *formal* basis we can show that the Georgian dative of time is not an indirect object, because the presence of such a dative is not formally correlated with indirect object markers in the verb form:

- (11.) ... oci čelia eg bemurazi *qovel sağamos* momadgebra...
'For twenty years this villain has been pestering me *every evening*...' [N. Dumbaje, *Me, bebia, Iliķo da Ilarioni.*]

In the above example, we cannot consider the dative *qovel sağamos* 'every evening' as an indirect object, since the verb form has no indirect object marking. This same reasoning can be applied to examples such as the following:

- (12.) a. Sakartveloši gavrclebuli endros pesvebi *ķarg šindispers* ijleva.. 'Madder roots, widespread in Georgia, give a *good cornell-cherry color*.' [Samšoblo No. 14 (735), 26 May 1989]

- b. Samokalako omis dros glexebi mtel *zedmeṭ p̄urs* mreṭvekobis, muṣebisa da ṭiteli armiis saṭiroebisatvis iṭleodnen. 'During the Civil War, the peasants gave all surplus grain for the needs of industry, workers, and the Red Army.' [KEGL s.v. *iṭleva*]

- (13.) *Gazetebi ert imistana ambavs iṭerebian*, romelic gverds ver aukcevs. 'The newspapers are writing *such a story* that one can't avoid it.' [KEGL s.v. *iṭereba*]

In these examples, all of which are limited to present series screeves only, the dative noun phrase is not correlated with any indirect object marker in the verb phrase; i.e., we do not find the forms *eṭleva* or *eṭereba* (both of which exist). Both of these verbs are generally agreed to be II. conjugation, hence, intransitive. If they are intransitive, then, they cannot have a direct object, a category limited to transitive verbs (I. and, less commonly, III. conjugation verbs). Thus, in these sentences, the dative noun phrase seems to behave somewhat like the dative of time in (11.) above.⁶

What is interesting about these examples is that they are basically active, and not passive in meaning. Thus, the grammatical subject in such constructions has a strong tendency toward being animate or personal. (In (13.), for example, the noun phrase *gazetebi* clearly has the meaning of 'newspaper writers.') The corresponding relative forms *eṭleva* and *eṭereba* are passive in meaning.

Closely related to this group are II. conjugation verbs which take two datives, one of which is marked in the verb form (and hence is to be regarded as an indirect object) and the other of which is not (and, thus, probably best not regarded as an indirect object). Examples:

- (14.) a. *Tvals naṭerṭṭali gadmomcivda*. 'Sparks fell from my eyes.'

⁶A similar construction with a dative that cannot be regarded as an indirect object is the dative of destination, now generally replaced by a postpositional phrase, e.g.:

Iciṭ, miṣṭer, me axlac *Svanets* miṭdivar... 'You know, Mister, I am going to Svanetia (dat.) now also.' [K. Gamsaxurdia, *Mtvaris moṭaceba*.]

In current usage *Svanet-s* would be replaced by *Svanet-ṣi*.

b. *Guls sevda šemomačva*. Grief grasped [lit. pushed against] my heart.'

c. *Tvalt damibnelda*. 'My eyes went blind.' [DK]⁷

d. *Mas kva daeca tavs* (tavze). 'A stone fell on his head.'

e. *Me bevri ucnauri ambavi gadamxdenia tavs*. 'A lot of strange things have happened to me.'

In the above and similar examples the subject is inanimate and the non-agreeing dative is a body part or *tavs* 'head' in an extended meaning. But examples also occur with both animate subjects and animate non-agreeing datives:

(15.) a. "Es *ķi mironis cxenia, švilebs gepicebi*." ambobs *Ĵoķia*..
 "This one is an outstanding horse, I swear to you on my children." *Ĵoķia* says.' [K. Gamsaxurdia, *Mtvaris moņaceba*]

b. *Sindiss gepicebi*. 'I swear to you on my conscience.' [KEGL s.v. *epiceba*]

c. *Ert rames tu dampirdebi*, ar mečqineba. 'If you promise me one thing, I won't be angry.' [DK]

In these examples, the subject is animate and the non-agreeing dative is animate (a.) or inanimate (b., c.). As in (14.) above, invariant is the animacy of the indirect object. The indirect object is, in general, always of higher relative animacy than the non-agreeing dative. This can be seen in the replacement of a non-agreeing animate dative in (15a.) with an inanimate in (15b.)

An example of an unmarked dative argument with a IV conjugation verb is found in (16.):

(16) *Txili čemis tvalis sinatlesac mırčevnia*. 'I (dat.) prefer a nut even to the light (dat.) of my eye[s].' [Važa Pšavela,

⁷Examples labeled *DK* were constructed by Professor Dodona Kiziria of Indiana University, to whom I would like to express my gratitude.

Čxiqyta korčili.]

Again, the agreeing dative is personal while the non-agreeing dative is not.

A third type of unmarked dative argument is found in I. conjugation (transitive) verbs. Examples:

(17.) *Sadils quradğeba mimikcie.* 'Pay attention to my dinner.'

[DK]

Such a sentence contains four actants, two of which (the second person subject and the first person indirect object) are formally marked by the verb form.

In (18.) the situation is somewhat more complex; the verb is I. conjugation, but has no direct object:

(18.) *Mağaziaši unda çavide da bavşvs mimixede.* 'I have to go to the store (and) [could you] look after my child/baby.'

[DK]

Further indication that the non-agreeing dative is not an indirect object can be found when these constructions are put into the perfect series. Here, as is known, the verb cannot mark what was the indirect object in the present/future and aorist series, and that noun phrase cannot appear as a dative. However, non-agreeing datives occur as datives in perfect series screeves:

(19.) a. *Aseti ubedureba mters tu dauțexia čvens tavs.* 'Only the enemy has brought such a misfortune on us.' [DK]

b. *Beds bevri gasaçiri daunatlia mis tavs.* 'Fate has blessed him with many hardships.' [DK]

c. *Kartvelebs mteri moulodnelad dasxmia tavs.* 'The Georgians unexpectedly attacked the enemy.' [DK]

In these perfect series constructions the non-agreeing dative functions almost exactly the same as a dative of time:

- (20.) *Çels es qaci ĵer ar minaxavs.* 'I haven't seen this man yet this year.'

Here the dative time expression remains in the dative in the perfect series of a first conjugation verb.

The "paradox" of perfect series indirect objects. One of the strangest features of the structure of Georgian is the fact that neither I. nor III. conjugation verb forms can mark the presence of an indirect object in the perfect series, while such a marking is possible for verbs of the II. conjugation (traditionally, "relative passives"). This is typologically unexpected, since the overwhelming majority of II. conjugation verbs are derived (denominal or derived from I. or III. conjugation verbs). Such II. conjugation verbs are in general more marked than I. or III. conjugation verbs and, consequently, we would expect the neutralization of oppositions in the marked perfect series to occur first in the more marked II. conjugation forms, and not in the I. conjugation verbs.

Connected with this problem is another problem: as I have tried to point out before, the presence of what have traditionally been regarded as indirect object markers of the 3rd person, in particular, in this instance, *e-*, don't always indicate the presence of an indirect object. Compare the following:

- (21.) a. *Masçavleblebma adreve icodnen, rom am dğes miğebuli nişani momaval çels simçipis aţestatşı çaeçerebodat.* 'The teachers knew in advance that the grade received this day would be entered into their [id.o.] high school diploma the following year.' (KEGL s.v. *çaeçereba*)

- b. *Raţom miliciaşı ar çaeçerebi?* 'Why don't you join the police (i.e., sign up for the police)?' (KEGL s.v. *çaeçereba*)

In (21a.) there is an indirect object, while in (21b.) there is none. Nevertheless, there is no indication of this difference in the verb forms in the present/future series or in the aorist series ((a.) *çaeçere*, (b.) *çaeçera*). However, in the perfect series there will be a formal distinction: (a.) relative *časçeria*, (b.) absolute: *çaçerila*. Again, the opposition relative/absolute in this and similar verbs is neutralized in all series

except the perfect, unlike in the I. conjugation, where the opposition relative/absolute is neutralized in the perfect series.

How can we explain these apparent exceptions to typological universals? I would like to propose the following: In Georgian, as a general tendency (though certainly far from absolute rule) indirect objects are animate or personal. There is a similar tendency for the subjects⁸ of I. conjugation verbs to be animate or personal, while the subjects of II. conjugation verbs tend to be inanimate. Therefore the typical I. conjugation verb in the perfect series usually has at least one animate or personal actant. Were the (generally animate) indirect object of relative II. conjugation verbs not to be marked, such verb forms would tend to have only inanimate actants in the perfect series. Thus, there appears to be a tendency in Georgian to have the verb form mark at least one animate actant whenever possible. In transitive verbs this actant tends to be the subject, hence the neutralization of the opposition absolute/relative in the marked perfect series. In II. conjugation verbs the animate actant tends to be the indirect object, hence the marking of the indirect object in the marked perfect series of these verbs.

Absolute II. conjugation verbs with "indirect object markers" tend to have animate subjects (see (21b.) above), so that in the perfect series an animate actant will already be marked by the verb form.

If this analysis is correct, it is further evidence in support of the central role played by the animacy hierarchy in the overall structure of Georgian, as pointed out by Tuite (1988).

Indirect objects and the animacy hierarchy. I would like to hazard one very preliminary hypothesis, which, I think, deserves further study.⁹ There seems to be a tendency for the "indirect object" to become the pivotal actant in the developing Georgian animacy hierarchy. In I. conjugation verbs the "indirect object" is *ideally* lower in animacy than the subject and greater in animacy than the direct object. In (derived) II. conjugation and in most IV. conjugation verbs, the "indirect

⁸In the following discussion, NP¹ will arbitrarily be called 'subject,' NP² – 'direct object,' and that noun phrase that is not marked in the perfect series of I. conjugation verbs or which is marked by *h*-series object markers in the perfect series of II. conjugation verbs will be called, again, arbitrarily, 'indirect object.'

⁹What follows does not, however, take into consideration the inversion in the perfect series of I. and III. conjugation verbs; a phenomenon, which for I. conjugation verbs at least is one of the oldest features of the language and certainly not an innovation

object" (dative) is *ideally* higher in the animacy hierarchy than the "subject" (nominative). The ideal indirect object (dative) is, generally a potential agent, unlike the ideal direct object.

Now, it is true that in many languages (in particular, the Indo-European languages of Europe) there is a strong tendency for indirect objects to be animate and direct objects to be inanimate.¹⁰ In fact, Georgian may appear to deviate from this tendency because of the widespread use of the indirect object to mark locative relationships of inanimate objects. (One should attempt to determine the productivity of these locative uses in contemporary literary Georgian.¹¹) Nonetheless the agreement patterns investigated by Tuite (1988) have shown the major role played by agreeing dative objects in the animacy hierarchy, which he has shown to be central to the understanding of the syntactic structure of Georgian.

Finally, we should probably view the tendency for the agreeing dative objects to be animate in conjunction with other innovations in Georgian tending to formally mark the animate/inanimate opposition. Let me mention only the strong tendency for the nominative plural marker *-ni* in pronouns (e.g., such forms as *qvelani* 'all,' *sxvani* 'others,' etc.) and numerals (e.g., *oriveni* 'both,' *samni* 'three (people),' etc.) to be restricted to animates, with the new nominative plural marker *-eb-i* tending toward being restricted to inanimates in pronominal forms such as *raebi*, *eseebi*, *egeebi*, etc.

REFERENCES

- Aronson, Howard I. 1979. Toward a typology of transitivity: the strange case of the Georgian subject. in: Paul R. Clyne, William F. Hanks, and Carol L. Hofbauer, eds. *The elements: a parassession on linguistic units and levels*, Chicago: Chicago Linguistic Society. Pp. 297-306.

¹⁰Note, for example, the formal neutralization of the opposition direct object/indirect object with that occurs with personal and animate nouns in Spanish. Compare:

Vi (unas) montañas. 'I saw some mountains.'

Vi a Juana. 'I saw Juana (direct object).'

Di a Juana el recado. 'I gave Juana (indirect object) the message.' (Green 1988:106-107)

¹¹Cf. the tendency to replace the dative marking of the superessive with a postpositional phrase.

- _____. 1991. Towards a typology of verbal categories. in: Linda R. Waugh and Stephen Rudy, eds. *New vistas in grammar: invariance and variation*. Amsterdam: John Benjamins. Pp. 111-131.
- _____. 1992. The marking of O₃ in Georgian.
- _____. Forthcoming. Paradigmatic and syntagmatic subject in Georgian. in: Howard I. Aronson, ed. *The non-Slavic languages of the USSR, linguistic studies. New series*. Chicago: Chicago Linguistic Society. Pp. 1-17.
- Filin, F. P., ed. 1979. *Russkij jazyk: ènciklopedija*. Moskva: Sovetskaja ènciklopedija.
- GDW = Tschenkéli, Kita. 1965-74. *Georgisch-deutsches Wörterbuch, 1-3*. Zürich: Amirani.
- Green, John N. 1988. Spanish. in: Martin Harris and Nigel Vincent, eds. *The Romance languages*. New York: Oxford University Press. Pp. 79-130.
- Harris, Alice C. 1981. *Georgian syntax: a study in relational grammar*. Cambridge: Cambridge University Press.
- KEGL = Čikobava, Arn., ed. 1950-64. *Kartuli enis ganmarṭebiti leksiṭoni, 1-8*. Tbilisi: Mecniereba
- Tschenkéli, Kita. 1958. *Einführung in die georgische Sprache, 1-2*. Zürich: Amirani.
- Tuite, Kevin. 1988. Number agreement and morphosyntactic orientation in the Kartvelian languages. University of Chicago Ph.D. dissertation.

MICROCOSMOS: THE CIRCASSIAN VERB

John Colarusso
McMaster University

1. Introductory Remarks. The Circassian¹ verb rivals the sentence in expressive power, apparently by making extensive use of a "syntactical" morphology that reflects both deep and surface syntactic structure. This verb is of great interest to contemporary linguistic theorists because it tells them what sort of information morphology must handle. To render such a grammatical marvel comprehensible to the reader it will be necessary to have a brief theoretical preamble. I have adopted autolexical syntax (Sadock 1991) in spirit without necessarily adhering to generalized phrase structure grammar (Sadock's preference) in technique. Autolexical syntax ('autolexical grammar' would be better) is a simple and yet powerful theory in which various grammatical components work in parallel, subject to simple constraints, to parse a given string in a variety of ways simultaneously. Thus a string's components have at one and the same time distinct linguistic significance, much as a note of music may enter simultaneously into melodic and harmonic structures. Symbolically one might represent such parallel processing as a Cartesian product of grammars, $G_1 \times G_2 \times \dots \times G_n$. Because these are rule systems acting upon strings the net result of such a product is a restricted set of acceptable grammatical forms (see Sadock 1991: 43-47, where a different representation of this process is given), rather than an expanded field of data as would be the result from a Cartesian product of sets.² Because the morphology

¹This paper is an elaboration of part of chapter four of Colarusso 1992. Circassian is a member of the Northwest Caucasian language family, which also includes the nearly extinct Ubykh, and Abkhaz and Abaza, which are closely related. Circassian is a cover term for Adyghe or West Circassian, a diverse complex of roughly seven dialects, and the distinct Kabardian (East Circassian), which is quite uniform. Besleney is transitional between Kabardian and West Circassian, but is closer in crucial respects to Kabardian than to any Western dialect.

²The Cartesian product of two sets, $A \times B$ of set A with elements $\{a\}$ and B with elements $\{b\}$, yields a new set with elements that consist of pairs $\{(a, b)\}$. Thus, if A is of size (cardinality) $\#(A) = a$, and B is of $\#(B) = b$, then the cardinality of $A \times B$ is $\#(A \times B) = a \times b$, which of course is bigger than the size of either A or B .

If, on the other hand, one is extending the notion of Cartesian product to that of grammars, then the opposite effect holds. Roughly, the probability of an element,

of a language seems to be one component with at most subcomponents in it, we shall see that despite several types of morphology acting in parallel to make up a verb, no element belongs to more than one subcomponent at a time. The Cartesian product of grammars does not hold, therefore, within a component (in other words the notion of component is still a coherent one) and instead one has a union of "sub-grammars" operating under complex informational and logical constraints in order to produce an intercalation of elements (Colarusso 1992).

Examples are from Bzhedukh West Circassian (unmarked) or Kabardian (K) East Circassian, I must thank Hisa Torkacho, Rashid Dahabsu, and Majda Hilmi for most of these forms.

2. Argument Structure. The Circassian verb wears its theta roles or argument structure (Sells 1985: 37-38; Haegeman 1991: 35-47, 286-294) on its sleeve. In (1) I have denoted the topmost node as either 'fulcrum' or 'absolutive.' This is the head of the argument structure in that the sense of the verbal action is determined by the noun in the absolutive for both transitive and intransitive verbs. That noun is the *fulcrum* upon which the verbal action is based and around which it pivots.

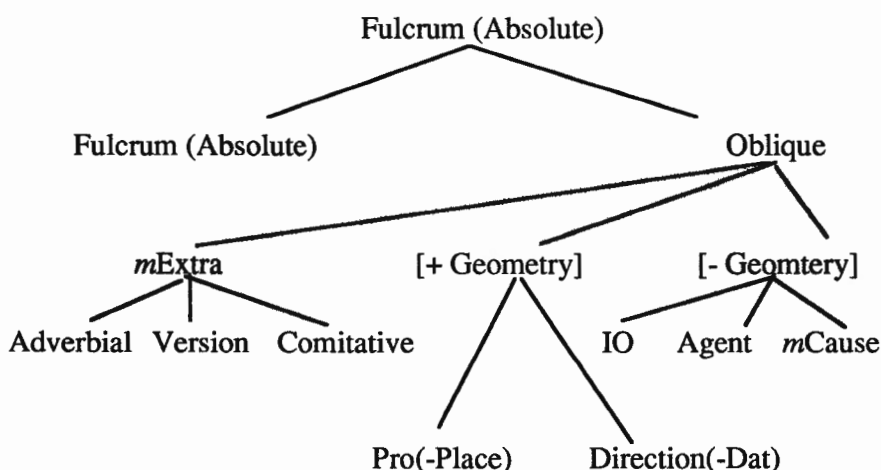
The fact that it is a head also implies that the other arguments will stand in an implicational hierarchy with regard to it. In other words the noun in the oblique, denoted by 'obl,' will imply that in the absolutive: obl → fulcrum (abs), and similarly for the other nodes. In this diagram

e, belonging in a derivation (tree or tress) of grammar G_1 ($p_e|_{G_1}$) is E^0/N^0 , where E^0 is the number of terminal nodes to which e may be assigned and N^0 is the number of terminal nodes in the tree(s) in question.. Similarly the probability of e being assigned to (a) tree(s) in G_2 is $p_e|_{G_2}$, and the probability that it will be in trees for both grammars simultaneously is $p_e|_{G_1} \times p_e|_{G_2} = p_e|_{G_1 \times G_2}$, which is less, of course, than being simply in one grammar.

This Cartesian product of grammars has the advantage of capturing the multiple role of an element while at the same time maintaining its unity and identity. In autolexical syntax as it now stands (Sadock 1991:20, 38-39, 43-47) the identity of an element between components is at best fortuitous and at worst illusory. This does not seem to be a desirable result. Also, my theory does not assume that the lexicons of the two grammars are identical (note Sadock's cautionary statement to this effect on page 224, note 16) since elements can exist that fall only within one grammar or the other and therefore do not constitute elements of the Cartesian product. Such elements would be morphemes of the pragmatical or derivational morphology.

and others to follow, unusual or special elements (perhaps derived from more complex structures), that is, nodes that are marked, will be marked by *m*. Thus, in (1) the causative node is marked to reflect the rarity, complexity, or difficult semantics of causative verbs. Such causatives may also be derived from a higher matrix verb, but arguments for this have yet to be found. I make the claim here that deep ergative languages, such as Circassian, are ergative precisely because the argument structures of their verbs has the headed structure shown (1).

(1) Argument structure



2.1. Fulcrum and Oblique [-Geometry] Indices. The ergative fulcrum is the initial index and refers to Ns in the absolutive. Other roles, considered traditionally to be grammatical (and hence [-geometry]), such as indirect object, agent (subject), or causative agent are non-initial indices and refer to Ns in the oblique case. To aid the reader I have underlined the relevant morphemes in this and the following sections.

It should be noted in all the following examples that once two or more arguments come between a prefix (of whatever type) and the stem that prefix shows a schwa-grade form. For example, in the strings prefix-io-s-stem, prefix-ag-cause-stem, or prefix-io-ag-cause-stem, the prefix will be in a schwa-grade. Contrast /-qa-/ 'horizon of interest' in

(2b, i) with /-qə-/ its schwa-grade in (2c, i, and ii). Also, phonological effects in the verb (a very rich area) are shown in parentheses to the right.³ For the abbreviations used in the examples, see the appendix.

(2) Fulcrum and oblique indices

a. Simple Intransitives

- (i) λ' ə-ha-r ma-a-p λ +a-ha (map λ á λ ˘ʰ)

man-pl-abs 3-pres-look+intr-pl

'The men are looking.'

- (ii) sa s-a-p λ +a (sáp λ a)

I I-pres-look+intr

'I am looking.'

b. Two-place transitives

- (i) λ' ə-m-a sa sə-qa-y-ha-ə- λ aY˘ə-Ya (səqah λ aY˘əY)

man-obl-pl me me-hor(izon of interest)-he-pl-non pr-see-past

'The men saw me.'

- (ii) sa wa wə-s- λ aY˘ə-Y (wəs λ aY˘əY)

I you you-I-see-past

'I saw you.'

c. Three-place transitives

- (i) a-š' wa sa sə-qə-w-a-y-ə-t^hə-Ya (səqwəyt^həY)

3-obl you me me-hor-you-to-he-non pr-give-past

³The complex phonology is hinted at in the parenthetical forms (see Smeets 1984). Most importantly /aa/ is realized as [aː] or [a] (Catford 1984), with the effect of spilling over from the syllable nucleus into the coda so that a following consonant does not color the preceding vowel, as would otherwise be the norm. It reduces to /a/ when unstressed. /ah/ and /afi/ have the same realization except that they do not reduce when unstressed. Nominal phonology is metrical (stress assignment is governed by feet), whereas verbal is not and several "primary" stresses can accumulate at the end of a verb. In Bzhedukh alone the sequence of /-s-/ followed by a stem initial coronal fricative (for example /š/) can result in an affricate (for example /č/). This nice example of an autosegmental process, while considered to be "real Bzhedukh," was not used by either of my two Bzhedukh friends, so I have not recorded it. In the glosses I have used either 'he' or 'it' to render the third person forms which are gender-blind in Circassian. No antifeminist posture on my part should be assumed from this. As to the plethora of sounds involved, the segmental inventory of Bzhedukh and the simpler Kabardian are shown in Appendix 2.:

'He gave me to you.'

- (ii) wa sa a-r \emptyset -qə-s-a-w-ŕ'a-a-ya (qəsapŕ'áay)
 you me 3-abs 3-hor-me-to-you-say-th v-past
 'You said it to me.'

d. Causatives

- (i) sa wa wə-s-ya-a-pŕ'a-a-ya (wəzɣapŕ'áay)
 I you you-I-cause-con-look-intr-th v-past
 'I made you look.'
- (ii) sa a-š' wa wə-qa-y-ə-s-ya-a-ŕay'ə-ya (wəqəyzɣaŕay'əy)
 or: sa a-š' wa wə-qa-y-a-s-ya-a-ŕay'ə-ya
 (wəqəyzɣaŕay'əy)
 I 3-obl you you-hor-3-non pr- (or -dat)-I-cause-con-see-past
 'I showed you to him.'
- (iii) sa a-ha-m-a wa a-r \emptyset -qə-w-a-y-ha-s-ya-a-t'ə-ya
 (qwahzɣat'əy)
 I 3-pl-obl-pl you 3-abs 3-hor-you-to-3-pl-I-cause-con-give-past
 'I made them give it to you.'

In many of the forms of (2) both pronouns and their corresponding verbal indices (affixes) can occur, though a rule of pro-drop often applies in discourse. Nevertheless, it seems that these indices cannot be viewed as mere clitics that have been incorporated into the verb. Nor can they be seen as echo pronouns ("You, (you) come here!"), since only one echo pronoun can occur in a sentence and by contrast there is no limit to the number of pronouns that can persist in the forms depicted in (2).

2.2. [+Geometric] Oblique Indices. These provide information about the physical or metaphorical stage upon which the action takes place. In (3c) a discontinuous stem is set off in brackets. The third person indices are built upon /-ə-/, unlike those of the indirect object.

(3) Geometric

- a. sa psə-ŕ'ə-š'a-m sə- \emptyset -ŕ'ə-a-ŕ (səŕ'áŕ)
 I river-edge-skin-obl I-3-edge-pres-lie
 'I lay on the river bank.'
- b. tɬə-ŕ-a-r ʔaana-m \emptyset - \emptyset -t'ə-y-a-a-ŕ (t'ayŕ)
 write-lic-th v-abs table-obl 3-3-surface-dir-dat-pres-lie

'The book lies on the table.'

- c. $\acute{s}^w\acute{a}$ -za-bğə-rə-[$\acute{x}a$]-t- γa -a-[$\acute{c}^w\acute{a}$]- γa ($\acute{s}^w\acute{a}$ zabğə $\acute{r}a$ $\acute{x}a$ d γa $\acute{c}^w\acute{a}$)
 y'all-recip-mid-dist-mass-we-cause-con-exit-past
 'We made y'all scatter.'

2.3 Extras These are a benefactive (or a detrimental), a comitative, and an adverb (often with instrumental force). They are not an obligatory part of the argument structure of any verb, but rather are merely extra referents which can be reflected in the verb.

The occurrence of a pronoun with a postposition or a separate word precludes use of a coreferential verbal index. Therefore, these are true clitics. With the benefactive, for example, there exists a separate postposition /p^haap'č'a/ (Kabardian /paap's'a/), which can be used with a pronoun. Similarly with the detrimental, adversative, comitative, and adverbial arguments separate postpositions or words exist and the verbal indices cannot then be used. Thus while obligatory arguments seem to trigger their indices in the verb, non-obligatory ones seem to rely upon the distinct mechanism of cliticization or incorporation to appear in the verb, note (4e). The version node characteristically takes an overt third person index, /-yə-/ singular, /-y-ha-/ plural. As a sign of its cliticization this index often does not undergo glide-vowel metathesis as would other overt third person indices (4b).

(4) Version

- a. p $\acute{s}a\acute{a}sa$ -r \emptyset -qə- $\acute{s}^w\acute{s}^w$ - \acute{a} -y-a-s- $\acute{s}^{hy}\acute{a}$ - γa (q $\acute{s}^w\acute{s}^w$ \acute{a} y $\acute{s}^{hy}\acute{a}$)
 girl-abs 3-hor-y'all-despite-dir-dat-I-lead off-past
 'I led the girl off despite y'all.' (said of marriage by abduction)
- b. p $\acute{s}a\acute{a}sa$ -r \emptyset -qə-y-ha- $\acute{s}^w\acute{a}$ -w-də-y-a-s- $\acute{s}^{hy}\acute{a}$ - γa
 (qy $\acute{a}hs^w$ \acute{a} bday $\acute{s}^{hy}\acute{a}$)
 girl-abs 3-hor-3-pl-despite-you-with-dir-dat-I-lead off-past
 'I led the girl off with your help despite them.'
- c. p $\acute{s}a\acute{a}sa$ -r wə-p^haap'č'a \emptyset -qə-y-a-s- $\acute{s}^{hy}\acute{a}$ - γa (qay $\acute{s}^{hy}\acute{a}$)
 girl-abs your-sake 3-hor-dir-dat-I-lead off-past
 'I led off the girl for you.'
- d. p $\acute{s}a\acute{a}sa$ -r \emptyset -qə-w-fə-y-a-s- $\acute{s}^{hy}\acute{a}$ - γa (qəpfay $\acute{s}^{hy}\acute{a}$)
 girl-abs 3-hor-you-for-dir-dat-I-lead off-past
 'I led off the girl for you.'
- e. *p $\acute{s}a\acute{a}sa$ -r wə-p^haap'č'a \emptyset -qə-w-fə-y-a-s- $\acute{s}^{hy}\acute{a}$ - γa (qəpfay $\acute{s}^{hy}\acute{a}$)

girl-abs your-sake 3-hor-you-for-dir-dat-I-lead off-past
 'I led off the girl for you.'

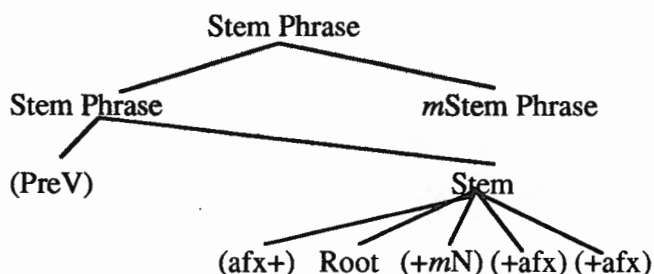
The adverbial (instrumental) node contains affixes that relegate the entire sentence to one of a sentential adverb. These may be translated into English as sentences with 'how,' 'when,' or 'why.' The verbal indices are copied, not incorporated, because they co-exist with an independent adverb (44a(i)). The discontinuous nature of 'when' /-zə-... -a-/ (5c) is remarkable and must be due to a morphological transformation (see §7.4).

(5) Adverbial index (embedding, control of S¹)

- a. yə-λas ø-x^wə-ya-x̂ tə-z-a-r-a-za-r-a-mə-λay^wə-ya-ra ø (x̂^wəyáx̂
 təzarazaraməλay^wəyara)
 gen-year 3-pass-past-already we-how-dat-instr-dat-recip-loc-
 dat-not-see-past-part copula
 'A year has already passed without our seeing one another.'
- b. ø-qə-z-fə-z-a-t^h-y-a-wəc^wa-a-ya-g^wara-ra-r
 (qəzfəzat^haywəc^wáayag^wararar)
 3-incept-what-for-self-dat-surf-dir-dat-stop-th v-past-
 referential-gerund-def
 'the particular reason why he stopped'
- c. sə-qə-zə-yə-yə-a-q^wə-ya-ž^wa sə-y-a-ž^wə-ya (səqəzərayq^wəyāž^wa
 sayž^wəy)
 I-hor-when-container-dir-when-be full-past-inst I-dir-dat-go-
 past
 'When I was satisfied I left.'

3. Stem Formation (Derivational Morphology). The verb stems that determine the argument structure show a wide range of derivational morphology. There is suffixation and prefixation, but there is also context-sensitive circumfixation, transparent preverb—root combinations, and even rare instances of verbal compounding (marked and somewhat poetic in tone). This derivational subcomponent is diagrammed in (6). Nodes within parentheses are optional. In (6) N = noun, Pre-V = preverb, *m* = marked, and *afx* = affix

(6) Derivational morphology



3.1. Suffixation. This is an important process. Four of the six suffixes are highly productive. Rarely two suffixes occur (7d). In some instances an incorporated noun (whence *mN* of (6)) comes between the root and two suffixes.

(7) Suffixation

- a. \emptyset -w-a-s- $\text{ʔ}^w\text{a}+\text{t}^h\text{a}$ -a- ʔa (was $\text{ʔ}^w\text{at}^h\text{áa}\text{ʔ}$)
3-you-dat-I-say+prolonged-th v-past
'I told it to you.'
- b. sa šxə -n-ə-m sə-y-a- $\text{šx}+\text{a}$ -a- ʔa (say $\text{šxáa}\text{ʔ}$)
I eat-inf-obl I-3-dat-eat+intr-th v-past
'I tried to eat the food.'
- c. sahbəy-a-r \emptyset -qə-w- ʔa -č' 'a -s- $\text{ʔ}+\text{h}\text{a}$ -a- ʔa (qəp $\text{ʔač}'\text{'as}\text{ʔh}\text{áa}\text{ʔ}$)
infant-th v-abs 3-hor-your-arm-instr-I-lie+active-th v-past
'I set the baby in your arms.'
- d. K yənə- zə -r \emptyset -zə- xa -t- ʔa - wək' $\text{'ə}+\text{r}\text{ə}+\text{y}\text{a}$ -śay-ś
(zə $\text{x}\text{ad}\text{y}\text{awək}'\text{'ə}\text{r}\text{əy}\text{á}\text{f}\text{ís}$)
giant-evil-abs 3-all-mass-we-cause-fall+smooth+down-past-aff
'We made the evil giant topple over.'

3.2. Prefixation. Prefixes can be familiar morphemes that appear elsewhere in other functions, as in (8a), or they can be forms that while transparent are nevertheless restricted in use, as in (8b).

(8) Prefixes

- a. tə-za-za+a+w-a-a-γa (təzazawáaγ)
we- recip- recip+con+hit-dat-th v-past
'We fought with each other.'
- b. sə-w-a-g^wə+p^hš^həsa-a-γa (səwag^wəp^hš^həsáaγ)
I-you-dat-heart+(tell a) tale-th v-past
'I was thinking about you.'
- c. zə-y-a-wə+fa (zaywófa)
self-he-pres-valence+bend
'He is bending over.'

3.3. Circumfixation. The occurrence of both a prefix and suffix is called a circumfix. Since one requires the other this is an instance of context-sensitive morphology (9). The suffix which converts a stative root into an active stem, /-hə-/ in (9a) is the same verb root as /-hə-/ 'to enter.' Form (9c) achieves its circumfixation by a context sensitive use of incorporation of a stem final noun, /dəyə/ 'joint.'

(9) Circumfixation

- a. sə-qa-t'ə+s+hə-a-γa (səqat'əsháaγ)
I-incept-down+sit+active-th v-past
'I sat down'
- b. tə-t^h-y-a-γ^wa+š^w+ə-γa (tə^hayγ^waš^wəγ)
we-surf-dir-dat-purpose+lie+dynamic-past
'We lay down (on a surface).'
- c. sə-qa-wə+q^wə+dəyə-γa (səqawəq^wədəyəγ)
I-incept-multiple/valence+stretch+joint-past
'I stretched/flexed myself.'

3.4. Preverbs. As in most languages basic verbal action in Kabardian can be modified by preposition-like particles affixed to the verb. Unlike many languages, however, the geometric sense of such "preverbs" is almost invariably maintained. I have placed these discontinuous elements in brackets. As (10c) illustrates, these can occur with the more ordinary geometric forms of the argument morphology. These derivational preverbs differ from the fulcrum or argument structure ones in that the former have no personal index referent whereas the

geometric morphemes of the argument structure can (§2.2, (3)).

(10) Preverbs

- a. ps-a-r ø-qə-[ḫa-čə]-ya (qəḫačəy)
water-th v-abs 3-hor-mass-flow out-past
'The water flowed out(as from a mass in a bucket or pot).'
- b. ps-a-r ø-qə-[də-čə]-ya (qədəčəy)
water-th v-abs 3-hor-opening-flow out-past
'The water flowed out (as from a hole or breach).'
- c. š'ə-za-bğə-rə-[ḫa]-t-ya-a-[č'ə]-ya (š'əzabğə-rəḫačəy)
y'all-recipr-mid-dist-mass-we-cause-con-exit-past
'We made y'all scatter.' (same as 3c)

3.5. Verb Compounding. Verb compounding is rare. I have only found instances where the verb /-č'ə-/ 'to exit, leave,' with or without a preverb, has been compounded with another stem.

(11) Verb compounding

- a. sa a-ha-m-a sə-ø-ha-[₁b₁la]-[₂p₂la]+[₁č'ə]-ya
(sahb₁ap₂lačəy)
I 3-pl-obl-pl I-3-pl-past (geometric sense)-look+exit-past
'I looked past them.'
- b. ø-za-bğə-rə-[y'a+ḫ+ə]+[č'ə]-ya-ha (zabğə-rəy'aḫəčəy^h)
they-recip-mid-dist-purpose+lie+dynamic+exit-past-pl
'They scattered and took up positions lying down.'
- c. K šə-r psə-m ø-ø-š'-a-[t'ə+s+ə]+[k'ə]-n-w-š'/
(š'at'əsək'yənəwš)
horse-abs water-obl 3-3-under-dat-[down+reach+dyn]+[exit]-
fut-def-aff
'The horse will drown in the water (river).'

4. Post-Stem Clitics. The suffixes following the stem are clitics⁴

⁴I call these terms clitics with some reservation. There seem to be three criteria for clitics currently in use. The oldest, (A), goes back to the original use of the term to denote a word that sometimes appeared as an affix. In this and the following criteria the order of the elements is immaterial. Many languages exhibit such clitics, including French, Ancient Greek, and the West Flemish dialect of Veurne (Sadock 1991:59, quoting Smessaert (1988)).

(A) Detachability: $X-i \longleftrightarrow X i$.

that attach to various parts of the syntactic tree (12). V^0 is the simple verb. It is the same as the stem phrase. V^1 is a verb-like node between the simple verb and the verb phrase. V^2 is the old verb phrase node (compare McCawley 1988:55ff.). INFL is short for 'inflection' and represents the various moods and tenses that can modify the verb phrase. S is the sentence node which is itself subject to adverbial and conjunctive modification. I shall present examples of the various elements going from left to right in (12). In (12) *pot* = potential, *neg* = negative, and *conj* = conjunction. Items within parentheses occur only in one sub-branch of Circassian.

The clitics seem to fall into natural groups modifying V^* by virtue of being headed on the right by adverb nodes. For reasons of space I have not gone into the detailed deletion arguments that would justify positing these internal heads, but have left this matter for future investigation. At present I have simply settled for semantic plausibility to motivate this presentation.

(12) Clitic train of inflectional morphology (herein C = complementizer)

One can have detachability with suppletive complications in the clitic, so that the clitic does not look exactly like its independent form, (B). This is actually fairly common, and would seem to play a role in some of the Circassian suffixes.

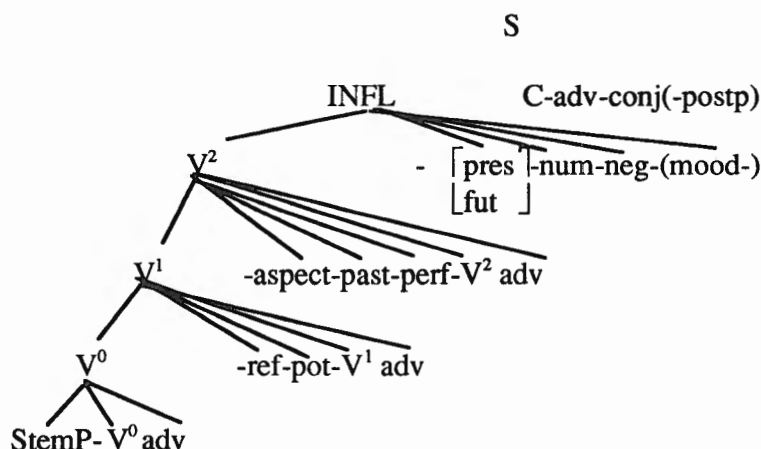
(B) Detachability with suppletion: $X-j \longleftrightarrow X-i$

A newer criterion is that of shifting affixation, wherein a set of affixes can appear on a different word if it is inserted in a sentence, (C). The best example of this is Bella Coola, wherein verbal inflection can appear on the negative particle and not on the usual verb, (D).

(C) Shifting affixation: $Y X-i \longleftrightarrow Y-i X$

(D) Bella Coola shifting affixation (after Davis and Saunders 1980: 44, 51, with slight changes in font).

- (i) Positive: ʔay-uc-m-tmaʔ*-a-k*-i-c'i-k t'aʔ*
do-mouth-mediopassive-reciprocal-they-quotative-contrastive-
perfective-unglossed those
'They were talking to each other while ...'
- (ii) Negative: ʔaʔ*-k*-i-lu-c'i-k ʔiiʔ*-anm-aw t'aʔ*
not-quotative-contrastive-expectative-perfective-unglossed far-limited
control development-they those



4.1. Stem Phrase Adverbs. Stem phrase adverbs or V^0 adverbs are a highly productive set of suffixes with a geometric sense (see Smeets 1983). These can occur with stem affixes to yield two adverb affixes in a row (13a) and can come with dative affixes of their own (13c). They often require particular preverbs and can hence exhibit context sensitive morphology.

(13) Stem phrase (V^0) adverbs

'They had not gotten far when ...'

The third criterion is that of exclusion: one cannot have a word and its clitic form co-existing in the same phrase, (E).

(E) Exclusion: *X x-Z

This criterion can also occur with suppletion. It is the criterion that I used to argue that most of argument morphology was non-cliticizing. The fourth criterion that one might put forward would be that of scrambling within a string of affixes, (F).

(F) Affix scrambling: X-i-j \longleftrightarrow X-j-i

To assume that (F) implies cliticization, however, is to beg the question regarding the nature and power of morphological grammars. It is possible that transformational shuffling of affixes, as we shall see in §7.4, is entirely within the scope morphology. (F) would only seem to imply cliticization if the scrambling of affixes had a correlate in the scrambling of syntactical variants, for here the Exclusion criterion would have to hold. We shall see that it does not hold for index scrambling in Kabardian (44,c).

Sadock (1991:61-62) also discusses two constraints on clitics and the configurational outcome of those constraints (page 69).

- a. sə-da-k'w+a+t^ha-yə-ž'ə-ya (sədak'w at^hayəž'əy)
I-verticle-move+intr+prolonged-up-reference-past
'I improved again.'
- b. tə-w-t^h-y-a-psaλə-ah-s'ə-ya (təpt^haypsaλahs'əy)
we-you-surf-dir-dat-talk-around-able-th v-past
'We were able to talk all about you.'
- c. yənə-z'-a-r q'əsħa-λaağa-m ø-y-a-ʔ'ə-λ'-a-a-ya (yaʔ'əλ'áay)
giant-evil-th v-abs mountain-high-obl 3-3-dat-stick-near-dat-th
v-past
'The evil giant was stuck (nailed) to the high mountain.'

4.2. Reference, Self, and Potential. Next come two morphemes with distinct roles. The first, reference (14a), is glossed as 'again.' It can express the fact that the verb implicitly assumes some earlier instance of the action or state. A form for 'self,' (14b) is used in intransitives that have a reflexive sense or, analogously, in three place transitives (or causatives of two-place transitives (14c)) to disambiguate the reflexive verbal index /zə-/. In such cases the preferred hierarchy of reference is first person before second and second before third. Reflexives without /-ž'ə-/ reverse the hierarchy (14d). Form (14b) is after Dumézil and Namitok (1939:26) and is in an old form of Chemgwi.

(14) Reference

- a. ø-qə-ø-t^h-y-a-y-ha-xə-ž'ə-ya (qə t^hərahxəž'əy)
3-hor-3-surf-dir-dat-3-pl-take-reference-past
'They took it back from him.'
- b. ø-qa-wəs'ə-ž'ə-ya (qawəs'əž'əy)
3-incept-awaken-self-past
'He woke up.'
- c. ʔ'əž'a-m-ž'a a-s' sa zə-sə-y-ə-ya-a-λay'ə-ž'ə-ya
(zəsəyyaλay'əž'əy)
mirror-obl-inst 3-obl I self-I-3-non pr-cause-con-see-self-past
'He let me see myself with a mirror.' or: *'He let me see himself with a mirror.'
- d. ʔ'əž'a-m-ž'a a-s' sa zə-sə-y-ə-ya-λay'ə-ya (zəsəyyaλay'əy)
mirror-obl-inst 3-obl I self-I-3-non pr-cause-see-past
'He let me see himself with a mirror.' or: *'He let me see myself with a mirror.'

As far as I can determine only Bzhedukh preserves a distinct form (15) with a final or 'at last' sense. The action or state referred to can be an intended or anticipated one. In all other forms of Circassian this has fallen together with that used for reference.

(15) Finality (Hadaghat'la 1968, vol. 1, p. 126)

psaása-r ø-q-ø-ha-s'wə-yə-y-ə-ħə-ž'ə-a-ya (qahs'wəryħəž'áay)
 girl-abs 3-hor-3-pl-despite-dir-3-non pres-carry off-final-th
 v-past
 'He carried off the girl despite them.'

Potential (16), is best rendered as 'can.'

(16) Potential

ø-qə-ø-t^h-y-a-y-ha-xə-ž'ə-s'wə-ya (qəthərahxəž'əs'əy)
 3-hor-3-surf-dir-dat-3-pl-take-reference-potential-past
 'They were able to take it back from him.'

4.3. V¹ Adverb. After reference and potential comes an adverb that seems to modify a verb node intermediate between the simple verb stem and what was traditionally called a verb phrase node.

(17) V¹ adverbs

- a. ø-y-a-šxə+a-ž'ə-s'wə-p^ha-ya-ha(yaşxəž'əs'əp^háyáx^h)
 3-3-dat-eat+intr-again-able-utmost-past-pl
 'They were able to feast again on it.'
- b. ø-s-šxə-s'wə-q^wa-ya-p(sšxəs'wəq^wáyáp)
 3-I-eat-able-excess-past-not
 'I was not able to eat too much.'
- c. ø-w-ŕ^wa-g^wá-n(pŕ^wag^wán)
 3-you-say-still-subjunctive
 'You would still say that?!'

Some of these adverbs (17a, b) have emphatic variants that are quite rare and have pragmatic senses. They might best be treated semantically as rare pragmatic suffixes (see §6), but as yet I have simply kept them with the V⁰ adverbs. These convey the judgment or attitude of the

speaker toward a situation. They exhibit a free floating lengthened emphatic stressed vowel either on the suffix itself, if the primary stress comes after it (18a), or otherwise a primary stress (18b). They are not context sensitive with any preverbs. Example (17a) is from Hadaghat'la (1968-1:121), (18b) from the same (p. 88), while (18b) is from Mr. Dahabsu.

(18) Emphatic (Pragmatic) V¹ Adverbs

- a. a-š' š^hə-ha-r ø-y-ə-ya-k'wə+a+də-p^ha-a-ya-š^{hy}t^hə-n
(əyak'wəadəp^haaγaš^{hy}t^hən)

3-obl horse-pl-abs 3-3-non pres-cs-go+intr+out-surely-emph-
past-perf-subjunctive

'Surely he has let the horses run off first thing!'

- b. ʒ'awaap^ha sə-z-a-r-a-š'wə-n ø-š^{hy}ə-ʔa-a-q^wa-ma (š^{hy}əʔaaq^wama)
answer I-how-dat-instr-dat-happen-inf 3-deix-exist-emph-
suffice-if

'Is it at all possible for me to help?'

In (18a) /-p^ha-/ seems to be derived from the same root as 'nose' or 'front,' with the semantic extension to 'first' as seen in /yah-p^ha-na-ra/ ordinal-first-ordinal-pronoun suffix = 'first,' while in (18b) /-q^wa-/ is obviously an adverbial use of the verb /(-y-a-)-q^wə-/ (direction-dative)-suffice = 'to suffice, fill up.' One should note the aberrant position of what appears to be the subjunctive /-n/ in (18a) as opposed to the more usual position in (19d). I have not found these emphatic variants in Kabardian.

4.4.Past Tense, Aspect, and Perfective. Numerous examples of the past tense /-ya/ have already been given. A double past is a remote past (19a), that is a past of a past.⁵ The past tense suffix can be preceded by a durative/habitual (19b) or progressive aspect, either with an assertive sense (19c) or a subjunctive one (19d), and followed by a perfective (19e). The perfective is clearly derived (at least

⁵Horizon of interest also is used to attest to the veracity of what is being said. A remote past, by definition falls beyond the sphere of personal knowledge and so normally cannot take a horizon prefix. The forms for verbs of atmospheric phenomena carry a special horizon preverb, and so constitute an exception to this rule of the distal past. The exact sense of the horizon preverb on weather verbs is not clear. See Colarusso 1984 for more details.

historically) from the auxiliary verb /šʰə-tʰ/ deixis-stand, which can be used periphrastically to express perfectives or statives (19f). In all other forms of West Circassian Bzhedukh /-tə-/ is replaced by /-šʰtʰə-/ (in Bzhedukh, Shapsegh, Natukhay, and Hakuchi) (/šʰtə/ in Chemgwi, Hatukhay, Yegerukhay, and Abadzakh). The durative or habitual past (19b) has a past tense suffix that shows an /-a/ where phonologically it should not. It is either an exception to the rule that deletes this vowel under most circumstances, or it is itself a small morpheme. I have glossed the form using the latter assumption.

Since the following adverb clearly modifies V² it seems that the past tense suffix does too. This sets it in contrast, however, with the present (and in Kabardian the future) affix, which modifies INFL.

(19) Past tense, aspect, and perfective

- a. wa-šʰxə ø-qə-y-a-šʰxə-ya-a-ya (qayšʰxəyáay)
sky-rain 3-chg of st-dir-dat-rain-past-th v-past
'It rained long ago.'
- b. wa-šʰxə ø-qə-y-a-šʰxə-tə-y-a(qayšʰxətəyá)
sky-rain 3-chg of st-dir-dat-rain-durative/habitual-past-
attributive affix
'It rained for a while.' *or*: 'It used to rain.'
- c. wa-šʰxə ø-qə-y-a-šʰxə-šʰtʰə-ya(qayšʰxəšʰtʰəyá)
sky-rain 3-chg of st-dir-dat-rain-prog-past
'It was raining.'
- d. wa-šʰxə ø-qə-y-a-šʰxə-nə-ya(qayšʰxənəyá)
sky-rain 3-chg of st-dir-dat-rain-subjunctive-past
'It might have been raining.'
- e. wa-šʰxə ø-qə-y-a-šʰxə-šʰtʰ-ya(qayšʰxəšʰtʰyá)
sky-rain 3-chg of st-dir-dat-rain-past-perf
'It had rained.'
- f. wa-šʰxə ø-qə-y-a-šʰxə-ya-n ø-šʰtʰ-tʰ(qayšʰxəyán šʰtʰ)
sky-rain 3-chg of st-dir-dat-rain-past-inf 3-deix-stand
'It had rained.'

The Kabardian system is simpler (20) (Colarusso 1992).

(20) Kabardian past statives and duratives

- a. Stative past

K λ' ə-z'-m psa λ a-n \emptyset - \emptyset -f'af'-ta (f'af't)

man-old-obl talk-inf 3-3-like-stand

'The old man liked to talk.'

b. Durative past (colloquial)

K sə- λ aaza-r-ta (sə λ azárt)

I-work-dist-stand

'I was working.'

c. Durative (aorist) past (formal or poetic language)

K sə- λ aaza-a γ -t (sə λ azáfit)

I-work-past-stand

'I was working.' or

'during the time that I was working'

Without past tense the aspect suffixes of Bzhedukh yield two futures, a general (21a) and an immediate (21b), which often has subjunctive or volitional force.

(21) Bzhedukh Futures

a. \emptyset -š'xə-tə (š'xət)

3-laugh-durative

'He will laugh.'

b. \emptyset -š'xə-nə (š'xən)

3-laugh-subjunctive

'He might laugh.' or

'He will laugh soon.' or: 'He shall (willingly) laugh.'

4.5. V² Adverbs After the past tense and its aspects comes another set of adverbs, apparently applying to V². These have a wide semantic range, but generally seem to set limits in time or quantity.

(22) V² adverbs

a. \emptyset -k'wə-a- γ a-a- γ a-xə-ha (k'wə γ a γ axəhá)

3-move+intr-past-th v-past-already-pl-aff

'They already left long ago.'

b. \emptyset -s-s'ə-s'ə- γ a-č'a-ha (ss'əs'ə γ ač'áx)

3-I-do-able-past-exhaustive-pl

'I was able to do them (the tasks) completely.'

c. \emptyset -s-šxə- γ a-a-š^{hy}-a (sšxə γ áaš^{hy})

3-I-eat-past-emph-excess

'I ate too much.'

d. \emptyset -s-šxə-ya-ɾʷa (sšxəyaɾʷ)

3-I-eat-past-little

'I ate just a bit.'

4.6. The Inflection Node. This node resembles the old aux(iliary) one. As one might expect because of its semantics (it is an inflectional node not a Vⁿ), it is not bounded by an adverb, thus distinguishing it from the other nodes in (12). Future and (a rare) present tense, number, negation, and mood are daughters of this node. In Kabardian plural number comes before future tense (23a).

(23) Plural number

a. K \emptyset -qə- \emptyset -t-y-a-y-ə-xə-žə-fə-ha-n-w-q'əm

(qətəɾəyχəžəfəɲanəwq'əm)

3-hor-3-surf-dir-dat-3-non pr-take-reference-potential-pl-fut-def-not

'He will not be able to take them back from him.'

b. \emptyset -qə- \emptyset -t^h-y-a-y-ə-xə-ž'ə-ś'ə-tə-ha-p (qət^həɾəyχəž'əs'ətəx^háp)

3-hor-3-surf-dir-dat-3-non pr-take-reference-potential-fut-pl-not

'He will not be able to take them back from him.'

The relic present tense /-ra/ occurs optionally in third plurals in Kabardian and obligatorily in present tense forms which are suffixed in some way, such as with a gerund (18b) or with a negative in both Kabardian and Bzhedukh.

(24) Relic Present

a. K ma-a-k'wə+a-ha-ɾ(-q'əm) (máak'wəɲar(-q'əm))

3-pres-move+intr-pl-pres(-not)

'They are (not) going.'

b. ma-a-k'wə+a-ha-ra-p (mak'wəx^háráp)

3-pres-move+intr-pl-pres-not

'They are not going.'

Kabardian has a very rich system of future tenses and various moods.

There seems to be no Western Circassian equivalent to this.

(25) Future and Moods

- a. K \emptyset -k'wə+a-ha-n-w-s (k'wəhənwəs)
3-move+intr-pl-fut-def-aff
'They will be going.'
- b. K sə-k'wə+a-aγ-sara (sək'wəfəsára)
I-move+intr-past-optative
'I would like to have gone.'
- c. K wə-q'ə-s-da-mə-ʔa+pəq'wə-aγ-ta-ma sa \emptyset -q'ə-s-a-ħaħa-k'ə-n-ta (wəq'əzdaməʔapəq'wəfətitama sa q'əzaħaħak'ənta)
you-hor-me-with-not-hand+help-past-irrealis-if I 3-hor-me-dat-hard-exhaustive-fut-irrealis
'If you had not helped me, it would have been very hard for me.'
- d. K wa a-bə wə- \emptyset -sə-s-ħaaγ'-aγ-sara-t (wəsəsħaaγ'wəfəsarat)
you 3-obl you-3-deixis-I-see-past-optative-irrealis
'If only I had seen you there!'

Perhaps the only trace of the irrealis /-ta/ in West Circassian is the concessive, /-γ'atʰ/ (dialect /-γatʰ/) (Keraševa 1960: 1088-1089). This suffix always has an emphatic vowel before it and it always conveys a past sense with or without an overt past tense affix.

(26) West Circassian Concessive

- a. sə-k'wə+a-a-γ-atʰ (sək'wəáγ'atʰ)
I-come+intr-emph-even if
'Even if I came ...'
- b. \emptyset -k'wə+a-γa-a-γ-atʰ (k'wəáγáγ'atʰ)
3-come+intr-past-th v-even if
'Even if he came ...'

4.7. Complementisers, Sentential Adverbs and Conjunctions The S-node itself can be modified by complementizers, adverbs, and conjunctions. The latter two, with conjunctions being very adverb-like in their syntactic function, bound the node on the right, just as other adverbs have done for the other nodes. The parallel between the S-node and the other V-nodes is, therefore, very close and tempts one to

suggest that the S-node is mislabelled. Instead it might better be termed a "maximal V."

(27) Kabardian Complementizer, Sentential Adverb, and Conjunction

- a. K \emptyset -q'a-k'" ω +a-a γ -ta-ma (q'ak'" ω afitama)
3-hither-move+intr-past-irrealis-if
'If he had come '
- b. K \emptyset -q'a-k'" ω +a-a γ -ta-ma- ω y (q'ak'" ω afitam $\acute{\omega}$ y])
3-hither-move+intr-past-irrealis-if-even
'Even if he had come ...'
- c. K \emptyset -q'a-k'" ω +a-a γ -ta-ma- ω y- ω y (q'ak'" ω afitam $\acute{\omega}$ y])
3-hither-move+intr-past-irrealis-if-even-and
'And even if he had come ...'

The Bzhedukh equivalents to the forms in (27) lack the irrealis /-t^ha-/ cognate, but are otherwise of the same pattern. Also in (27) I have used only one complementizer, /-ma/, merely to indicate position. In (28) I give Bzhedukh examples of the others, using [...] to denote embeddings. The purpose complementizer (28g) is exceptional in that the verbs it governs seem not to be inflected for person even when overt nouns remain under their control.

(28) Complementizers

- a. Zero complementizer (doubly embedded)
[[sa z $\acute{\omega}$ -g ω ara \emptyset -q $\acute{\omega}$ -s-a- $\acute{\omega}$ -a] \emptyset - \emptyset -fa-da] \emptyset -q $\acute{\omega}$ -s-s' ω a-s' $\acute{\omega}$ - γ a
(q $\acute{\omega}$ s $\acute{\omega}$ a f $\acute{\omega}$ da q $\acute{\omega}$ s s' ω as' $\acute{\omega}$ Y)
me one-certain 3-hor-me-dat-call-dat 3-3-for-be like 3-hor-me-
despite-know-past
'It seemed to me as though some one called to me.'
- b. Infinitive complementizer
[$\acute{\omega}$ ' $\acute{\omega}$ -m s' ω z-a-r \emptyset -qa-y- $\acute{\omega}$ -b $\acute{\omega}$ na-n] \emptyset - $\acute{\omega}$ - γ a (q $\acute{\omega}$ y b $\acute{\omega}$ na n
 $\acute{\omega}$ ' $\acute{\omega}$ Y)
man-obl woman-th v-abs 3-incept-3-non pr-desert-inf 3-
happen-past
'It happened that the man deserted the woman.' or: 'It was possible that the man deserted the woman.'
- c. Predicative (declarative) complementizer
[$\acute{\omega}$ z $\acute{\omega}$ -qa-y- $\acute{\omega}$ -m $\acute{\omega}$ - $\acute{\omega}$ - $\acute{\omega}$ - $\acute{\omega}$] \emptyset -s' ω -t^h (z $\acute{\omega}$ q $\acute{\omega}$ y m $\acute{\omega}$ $\acute{\omega}$ - $\acute{\omega}$ - $\acute{\omega}$ s' ω -t^h)

self-incept-3-non pres-not-stand up-pred 3-deixis-stand
 'He should not get up.' (lit., 'That he not raise himself up
 stands there.')

- d. Infinitive-predicative (declarative) compound complementizer
 [sʰə-m sʰəz-a-r ø-qa-y-ə-bgəna-n-əw] ø-mə-šʰ
 (qəybğənanəw məšʰ)

man-obl woman-abs 3-incept-3-non pr-desert-inf-pred 3-not-
 happen

'It is impossible for the man to desert the woman.'

- e. Indirect discourse, special /-ma/ form

sa wa [sʰə-a-r ø-s-šə-ya-ma] ø-w-a-s-ʰa-a-ya (sʰəšəyama
 wasʰəay)

I you [meat-th v-abs 3-I-eat-past-that] 3-you-dat-I-say-th
 v-past

'I said to you that I ate the meat.'

- f. Indirect question complement as a definite (absolutive) nominal

[Nzə-z-fa-w-ya-wəcʰa-a-ya-ra-r] ø-s-a-sʰə
 (zəzfəyawəcʰəayarar səsʰə)

[self-what-for-you-cause-stop-th v-past-gerund-abs] 3-I-pres-
 know

'I know why you stopped.'

- g. S-adverb or specialised complementizer 'in order to'

[ʰhaasʰa-r (ø-ø)zʰa-nə-kʰa] ø-da-čʰə-ya (zʰanəčʰəa dačʰəy)
 [field-abs (3-3-)plough-inf-in order to] 3-out-exit-past

'He went out to plough the field.'

- h. Question complementizer

a-r wə-šʰə ø-ø-ra-ah (pʰərah)

3-abs your-brother 3-be-pres-Q

'Is he your brother?'

- i. Intensive/rude complementizer (compare /ba/ 'much')

a-r wə-šʰə ø-ø-ra-ba (pʰərabə, pʰəbə)

3-abs your-brother 3-be-pres-indeed

'He is indeed your brother!'

- j. Intensive/rude question complementizer

a-r wə-šʰə ø-ø-ra-ba-ah (pʰərabáh, pʰəbáh)

3-abs your-brother 3-be-pres-Q

'Is he indeed your brother?!'

4.8. Postpositional Sentential Adverb The West Circassian dialects

have an elaborate sentential adverbial postposition that is lacking from Kabardian and its cogener, Besleney (Paris 1974:232-233, §§228, 229). The form in (29) is in Shapsegh (after Paris 1974:44).

(29) Shapsegh Adverbial Postposition

ø-qa-a-k'wə+a-ra-psawə-m ø-y-ha-y-ə-γa-λay'ə-ra-əy-g'a-
za-γ'a-m (yahreyγaλay'əəyγ'azaγ'am)

who-hor-pres-come+intr-pres-all-obl it-3-pl-3-non pres-cause-
see-emph-and-still-all-time-adv

'He always shows it to all who come.'

4.9. Conjunction and Verb-Raising In West Circassian there is a consecutive conjunction /-s(a)/ 'and then,' which seems to take the place of both a complementizer and a conjunction. This form (30a) permits the unusual phenomenon of verb adjunction (30b). The resulting mammoth form must have equal subjects, but the subject indices show no signs of equal-subject-deletion. Such verb adjunction is distinct from the verb-raising that is seen when the object of an embedded verb is generic (30c and d). The Bzhedukh form (30b) is from Hadaghat'la (1968, vol. 1, p. 161).

(30) Bzhedukh conjunction affix, verb adjunction, and verb-raising

a. wə-ø-χa-wə+s'k'wə+ta-nə-sa wə-wə+s'k'wə+ta-ž'ə-n
(wəχawəs'k'wətanəs wəwəs'k'wətaž'ən)

you-3-mass(in)-valence+slip+down-fut-and then you-non
pr-valence+slip+down-self-fut

'You will slip down in the mud and then loose face.'

b. wə-ø-χa-wə+s'k'wə+ta-nə-s-w-ə-wə+s'k'wə+ta-ž'ə-n
(wəχawəs'k'wətanəswəwəs'k'wətaž'ən)

you-3-mass(in)-valence+slip+down-fut-and then-you-non
pr-valence+slip+down-self-fut

'You will slip down in the mud and thus loose face.'

c. sa λə ø-s-šxə-γa ø-s-s'wə+yəγ'a-a-γa (sšxəγa sš'wəyəγ'áay)

I meat 3-I-eat-past 3-I-subjectivizer+want-th v-past

'I wanted to eat meat.' (lit., 'I wanted it that I ate meat.')

d. sa λə ø-s-šxə-s-s'wə+yəγ'a-a-γa (λəsšxəss'wəyəγ'áay)

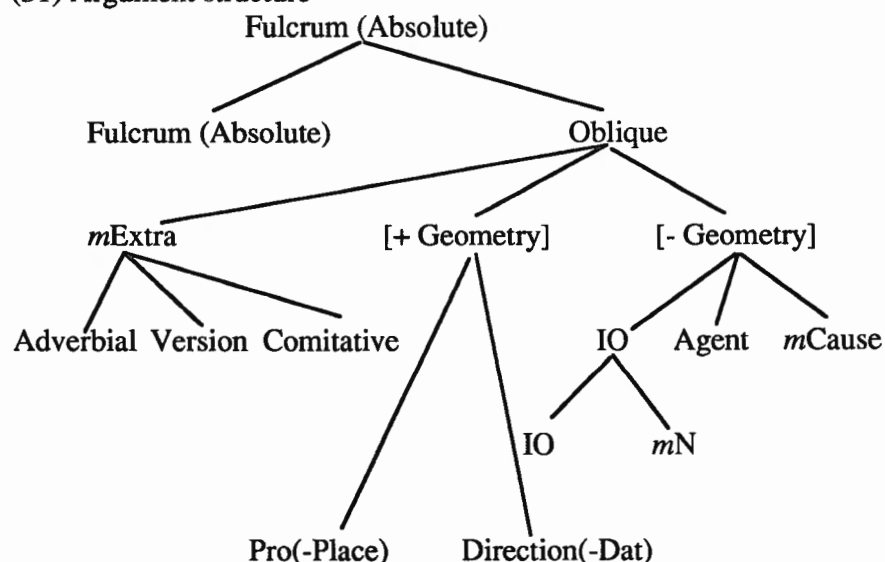
I meat-I-eat-I-subjectivizer+want-th v-past

'I wanted to eat meat.'

5. Incorporation. Three sites permit incorporation (Sadock 1991:78-110; Baker 1988a) of a noun.

5.1. Sites. Nouns may be incorporated as stem suffixes (*mN* of (6)), as preverbs with an instrumental sense, or as elements that can modify indirect-objects. For the last two functions (1) must be enlarged to that in (31). These incorporated nouns can take overt third person indices, the indirect objects whose direction or locus they can serve to elaborate, whereas preverbs take zero third person indices.

(31) Argument structure



5.2. Pre-Root Incorporation Pre-root incorporated nouns can show an overt indirect object verbal index (32a), and are usually generic, (32b, c, d). Note in (32b, c, d) that such incorporated Ns come after any preverbs in the verbal complex. The Kabardian form (32b) is from Hadaghat'la (1968 1:263). It exhibits /-yə-yə-/ container-direction (= in), an idiom meaning 'having the property of (being) ...'

(32) Pre-root incorporated N

a. sə-y-a-λa-ʔa-a-ya (sayλaʔáay)

I-3-dat-footpath (leg)-say-th v-past

'I asked/beseeched him.'

- b. K də-yə-yə-thə+ʔə-ḵ-ə+d-s (dəɾəy^həʔ^wəḵ^wəd^s)
we-cont-dir-god+front-bend+down-aff

'We are the ones who bow down before god. (a pagan worshipper)' (Hadaghat'la 1968, 1, §76: 263-5)

- c. K sə-ḵə-a-psaḵə-hə-n-w-s (səḵapsaḵəhənəw^s)
I-mass-dat-word-enter-fut-def-aff

'I shall enter the conversation.'

- d. K də-ø-yə-rə-yə-da+zə-a-w-k^wə+a (dəɾəydazáwk^wa)
we-3-groove (= path)-dist-dir-out+turn(= circuit)-pres-prog-move+intr

'We are going for a long walk.'

5.3. Stem Noun Incorporation in Adverb Position The stem adverb position (6) is already marked with *mN* to reflect the incorporation in (33). (Note circumfixation by incorporation in (9c).)

- (33) Stem phrase noun incorporation into adverb position

sə-qa-wə+q^wa+dəyə-ya/ (səqawəq^wadəyáy^y)

I-change of state-multiple+stretch+joint-past

'I stretched (flexed) myself.'

5.4. Instrumental The instrumental nouns that are incorporated (into the adverbial node under the Extras node) appear to occupy a position reserved for subordinating indices (5). Incorporated nouns in this position can not only be referential (as opposed to the claim that incorporated nouns must always be generic), but can even show possession.

- (34) Instrumental

wə-s+ʔa+č^ya-k^wə+a+də-n (wəsʔač^yak^wadən)

you-my+arm/hand+inst-move+intr+out-fut

'You will die by my hands.' *or*: 'You are about to die by my hands.'

5.5. Absolutive Fulcrum In an old form of Chemgwi used by the poet Tsey Ibrahim one finds an inalienably possessed noun incorporated instead of an index into the initial position of the verbal complex. It is

also marked with a detrimentive despite its absolute fulcrum position (Tsey Ibrahim, story II, line 16 in Dumézil and Namitok 1939:30).

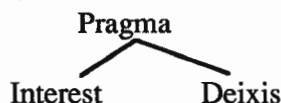
(35) Old Chemgwi

w-šh̄a ø-(ø-)s'wə-s-č'ə-na → p-šh̄a-s'wə-s-č'ə-n
 your-head 3-(3-)detriment-I-tear-volitional fut
 'I shall tear off your head.'

This phenomenon is regularly applied in Abkhaz and Abaza to neuter singular direct objects and plural direct objects (Ns reflected by /y-/ index) when they precede the verb.

6. Pragmatical Morphology. Pragmatical morphological affixes do not have an independent correlate anywhere in the sentence (see also Scancarelli 1986). Rather, they reflect the speaker's pragmatic stance with regard toward the affairs expressed by the utterance. These affixes occupy a position after the initial index (if overt) or the adverb index. The pragmatic affixes appear to be an important "signpost" in the trek across the verb. They are highly productive and of diverse significance. Their structure is that in (36). I have labelled the maximal node "pragma."

(36) Pragmatical Structure



6.1. Horizon of Interest Verbs reflect the speaker's (by polite extension, sometimes the hearer's) physical, social or emotional stake in the noun in the absolutive. Hence, it is an ergative morpheme. Verbs carrying this affix are termed /ʔa-k'wac'ə/ arm (hand)-inside = 'in the hand,' as opposed to those without, which are termed /ʔa-č'əb/ arm (hand)-back = 'outside the hand.' Their range of meanings is fascinating, progressing from a least marked personal sense of territory upon which an action takes place, to a marked expression of kinship involvement. This affix (37) is homophonous with a change of state or inceptive one (38) (see Colarusso 1984 for a full discussion). In its kinship sense (37d) /-qa-/ is context sensitive and requires an illative stem affix /-a-/.

(37) Horizon of interest

- a. sa wa t̃əʎ-a-r ø-w-a-s-t̃ə-ya (wast̃əʎ)
I you book-th v-abs 3-you-dat-I-give-past
'I gave you the book.'
- b. sa wa t̃əʎ-a-r ø-qə-w-a-s-t̃ə-ya (qwest̃əʎ)
I you book-th v-abs 3-hor-you-dat-I-give-past
'I loaned the book to you,' that is to say, 'I gave the book to you, but have retained an interest in it.'
- c. š̃ə+wə-m ʎ̃'ə-a-r ø-y-ə-wəč̃'ə-ya (əwəč̃'əʎ)
horse+man-obl man-th v-abs 3-3-pres-kill
'The horseman killed the man (a stranger).'
- d. š̃ə+wə-m ʎ̃'ə-a-r ø-qa-y-ə-wəč̃'ə+a-a-ya (qəwəč̃'əáy)
horse+man-obl man-th v-abs 3-hor-3-pres-kill+in-th v-past
'The horseman killed the man (speaker's kinsmen).'
- (38) Change of state (inceptive)
- a. ʎ̃'ə-a-r š̃ə-m ø-ø-ya-psə-χə-ya/ (yapsəχəʎ)
man-th v-abs horse-obl 3(man)-3(horse)-vertical-descend-down-past
'The man dismounted (but may soon remount).'
- b. ʎ̃'ə-a-r š̃ə-m ø-qa-ø-ya-psə-χə-ya (qaypsəχəʎ)
man-th v-abs horse-obl 3(man)-incept-3(horse)-vertical-descend-down-past
'The man dismounted (and is finished riding).'

6.2. Verbal Deixis The verb can also denote a particular place where the action took place, quite apart from any details regarding geometric information. This can govern a pronominal index (/ø-(ha-)/) which is reflected by a deictic pronoun in the syntax.

- (39) Deixis of verbal action
- a-š̃ sa č̃'aaʎa-m-a sə-ø-ṣ̌̃ə-y-ha-fa-zə-ya (səṣ̌̃(y)ahfazəʎ)
3-obl I child-obl-pl I-3-deixis-3-pl-for-turn-past
'I met the children there.'

6.3. Order. The order of the three nodes is demonstrated in (40), which though an awkward form can still be elicited.

- (40) Order of pragmatical affixes

sə-qə-z-fə-š^{hy}ə-y-ha-fa-zə-ya-ra (səqəzfəš^{hy}(y)ahfazəyar)
 I-hor-what-for-3-deixis-3-pl-for-turn-past-gerund
 'the reason why I met them there (on my territory)'

7. The Nature of the Verb. There are features exhibited by the Circassian verb which suggest that it is more than a word by the standards of English and yet less than a sentence. I shall briefly examine five such features: variable ordering and scope, differing roles, homophony, transformations, and the cross-over constraint. All of these, apply to whole words or sentences as well. Their occurrence at or below the word level is rare, but it strongly suggests that the morphological component is not formally different from the syntactical one, merely that it is under utilised.

7.1. Variable Ordering and Scope. Some clitic adverbs can occur in differing orders. This is to be expected from their sense and neutral semantic scope, as well as from their simultaneous role in syntax. Morphology here reflects syntactical freedom.

(41) Variability (differing orders, differing scopes)

- a. ø-wə-ʎa+č^{hy}ə-ya-ma-əy (pʎač^{hy}əyaməy)
 3-you-[unanalyzable preverb]+can-past-if-even
 'Even if you could do it.'
- b. Old Chemgwi (Tsey Ibrahim, p. 27, story I, l. 20, in Dumézil and Namitok 1939)
 wə-qa-pʎa+a-n ø-w-ʎak^{hy}ə-ya-ya-ma (pʎak^{hy}əyayama)
 you-hor-look+at-inf 3-you-can-past-only/even-if
 'If only you could look at (it).'
- c. K. ʈa-m ø-ž-y-ə-ʔa-ya-ma (žyʔáyma)
 god-obl 3-back-3-non pres-say-only/even-if
 'God willing!' (lit. 'If only God says it!')
- d. K də-ø-yə-rə-yə-da+zə-a-w-k^{hy}ə+a (dəɾəydazáwk^{hy}a)
 we-3-groove-dist-dir-out+turn(= circuit)-pres-prog-move+intr
 'We are going for a long walk.'
- e. tə-da+zə-y-a-k^{hy}ə+a-a-ya (tədazayk^{hy}áay)
 we-out+turn-dir-dat-move+intr-th v-past
 'We acted in that matter.'
- f. Shapsegh West Circassian (Hadaghat'la 1968-1:137-40, #17)
 ø-y-ə-s^{hy}ʔa-š^{hy}t^hə-ya-p-t^ha-əy (əs^{hy}ʔaš^{hy}t^həyapt^həy)

3-3-non pr-know-durative-past-not-stand-and
 'He was in a state where he had been ignorant of it and ...'

g. Shapsegh

ø-y-ə-s'ʔa-ya-t^ha-p-əy (əs'ʔayət^hapəy)

3-3-non pr-know-past-stative-not-and

'He had not known it.'

h. K sa wa a-r ø-q'ə-w-žə-s-ʔa-žə-aγ-s (q'əbžəsʔažáfis])

I you 3-abs 3-hor-you-again (= back at)-I-say-again-past-aff

'I said it to you again.'

The differing order of the subjunctive has already been noted in (18a) and (19d).

7.2. Differing Roles. The verbal morphemes show differing roles. These maintain their semantics pretty much unaltered, but change their behaviour with regard to position and index government.

(42) Differing roles

a. Regular horizon of interest: Shapsegh (Hadaghat'la 1968-1:137-40, #17)

ø-q-ø-ha-p^ha-n+a-n ø-š^{hy}ə-mə-ʔa-w (qáhp^hanan š^{hy}əməʔaw)

3-hor-3-pl-nose-remain+in-inf 3-deixis-not-be-pred

'There is no one to stand guard over them.'

b. Horizon as a preverb governing an index: Shapsegh (Hadaghat'la 1968-1:137-40, #17)

s-ø-ha-qə-p^ha-a-n+a-ma ø-x^wə-š^{hy}t^hə-ba-ah (sahqəp^háanama š^wəš^{hy}t^həbáh)

I-3-pl-hor-nose-dat-remain+in-if 3-happen-durative-much-question

'Of course isn't it possible if I stand guard over them at their place?'

7.3 Homophony Affixes, just like words, can show homophony.

(43) Homophones

a. K ø-z-a-p^hλ+a-k'ə-y^wa ø-y-ə-mə-š^wa-aγ-wa (zap^hλak'əy^wa yəməš^wáfíwa)

3-self-dat-look+at-behind (or) exhaustive-time 3-3-non pr-not-fit-past-pred

- (i) 'that he did not find the time to look behind himself'
- (ii) 'that he did not find the time to look himself all over'
- b. tə-z(-)a-r-a-ʎaɣʷə-ʒʷa (təzaraʎáɣʷəʒʷa)
we-recip (or) how(-)dat-instr-dat-see-inst
- (i) 'how we saw (something)'
- (ii) 'with our seeing one another'

7.4 Transformations Most dramatically the Circassian verb can show a limited amount of shuffling due to transformations. The most extensive amount is in the present tense, jussive, and negative of subordinated, optative, and irrealis forms, where these elements, normally expected to be suffixes, appear before the verb stem. Sometimes scrambling is accompanied by suppletion, as in (44b), the disability forms. One might also argue that the few inflectional affixes of jussive mood, present tense, and, for Kabardian, progressive aspect, that occur root initially (44d, e), are also preposed from the usual inflectional clitic train by a transformation.

(44) Morphological transformations

a. Negative flip in optative and irrealis

- (i) K wa a-bə wə-θ-sə-sə-mə-ʎaaɣʷ-ay-...-sara-t
(wəsəzəməʎaaɣʷafisarat)
you 3-obl you-3-deixis-I-not-see-past-...-optative-irrealis
'If only I had seen you there!'
- (ii) K wə-q'ə-s-da-mə-ʔapəqʷə-ay-...-ta-ma sa θ-q'ə-s-a-ħaħa-k'ə-n-ta (wəq'əzdaməʔapəqʷəafitama ... q'əzaħaħak'ənta)
you-hor-me-with-not-help-past-irrealis-if I 3-hor-me-dat-hard-exhaustive-fut-irrealis
'If you had not helped me, it would have been very hard for me.'

b. Disability

- (i) θ-qə-w-fə-θ-tʰ-y-a-s-ħə-žʷə-śʷə-ɣa-ha-p
(qəpfətʰaysħəžʷəśʷəɣáħʰáp)
3-hor-you-for-3-surf-dir-dat-I-take-back-able-th v-past-pl-not
'I was not able to take them back from him for you.'
- (ii) * θ-qə-w-fə-θ-tʰ-y-a-s-fa-ħə-žʷə-...-ɣa-ha-p
(qəpfətʰaysfaħəžʷəɣáħʰáp)
3-hor-you-for-3-surf-dir-dat-I-take-back-able-past-pl-not

'I was not able to take them back from him for you.'

(iii) \emptyset -qə-s-fə-w-fə- \emptyset -t^h-y-a-... \hat{x} ə-ž'ə- γ a-ha-p

(qəsfəpfət^hay \hat{x} əž'ə γ á \hat{x} ^háp)

3-hor-you-for-3-surf-dir-dat-I-take-back-able-past-pl-not

'I was not able to take them back from him for you.'

c. Contrastive scrambling (only in Kabardian; Kardanov 1955: 1035)

(i) K a-bə sa s'aa λ a-ha-m sə-q'ə- \emptyset -sə-y-ha- \hat{x} ^a-zə-a γ -s

(səq'ə γ ah \hat{x} ^aazáfis])

3-obl I child-pl-obl I-hor-3-deixis-3-pl-for-turn-past-affir

'I met the children there.'

(ii) K sa s'aa λ a-ha-m a-bə sə-q'ə-y-ha-sə-.....- \hat{x} ^a-zə-a γ -s

(səq'ahsə \hat{x} ^aazáas)

I child-pl-obl 3-obl I-hor-3-pl-deixis-...-for-turn-past-affir

'I met the children there.'

d. Jussive pre-stem affix

\emptyset -wara-k'^wə+a-... (warák'^wa)

3-jussive-come/go+intr-

'May he go!'

e. Present progressive

K s-a-w-k'^wə+a-...-... (sáwk'^wa)

I-pres-prog-come/go+intr-

'I am going (along).'

The scrambling in disability forms (44b) is sensitive to ergative forms, since it does not occur when the relevant index is an absolutive (that is coreferential with a nominal in the absolutive) (45).

(45) Blocking of disability scrambling

a. *sa t \hat{x} ə λ ə-m sə-fə-y-a-...-mə-ž'a (səfayməž'a)

I(abs) book-obl I-for-3-dat-not-read

'I cannot read the book.'

b. sa t \hat{x} ə λ ə-m sə-y-a-mə-ž'a-s'^wə (sayməž'aś'^wə)

I(abs) book-obl I-3-dat-not-read-able

'I cannot read the book.'

Reciprocals show both movement (46a) and replacement

transformations (46b). The exact derivation of these forms is unclear. They are clearly complex. Some are anti-transitives morphologically (46a) (traditionally termed “anti-passives” wherein the usual object becomes an indirect object and the force of the verb is weakened), even though the syntax is still ergative, while others show scrambled causative forms (46b). Clearly in the case of (46b), with its causative affix, simple copying of some scrambled pronouns from the syntax is not enough to account for the behaviour of the morphology. Whatever the exact nature of their underlying derivation, the morphology undergoes scrambling and suppletion in accord with the semantics of reciprocity. Therefore the morphology itself must permit transformations within its own component.

(46) Reciprocals

a. Two-place transitive

(i) Underlying order

zə-m zə-r za-qa-y-ha-ʃaγʷə-ya-ha

recip- obl recip- abs recip- hor- 3- pl- see- past- pl

(ii) Anti-transitive movement (scrambling?)

zə-r zə-m za-qa-y-ha-ʃaγʷə-ya-ha

recip- abs recip- obl recip- hor- 3- pl- see- past- pl

(iii) Surface form

ø-qə-za-r-a-ʃaγʷə-ya-ha (qəzaraʃaγʷəyáx̂)

3-hor-recip-inst-dat-see-past-pl

‘They saw one another.’

b. Causative of three-place transitive

(i) Underlying form

zə₁-m zə₂-m a-ha-r ø-qə-y-a-y-ə-tʰə-žʷə-ya-ha

recip-obl recip-obl 3-pl-abs 3-hor-3-dat-3-non pres-give-back-
past-pl

(ii) Scrambling (“anti-dative” movement, “causative agent” movement ?)

zə₂-m zə₁-m a-ha-r ø-qə-y-a-y-ə-tʰə-žʷə-ya-ha

recip-obl recip-obl 3-pl-abs 3-hor-3-dat-3-non pres-give-back-
past-pl

(iii) Surface form

zə₂-m zə₁-m a-ha-r ø-qə-y-ha-za-r-a-ya-a-tʰə-žʷə-ya-ha

(qahzarayətʰəžʷəyáx̂)

recip-obl recip-obl 3-pl-abs 3-hor-3-pl-recip-inst-dat-cause-
con-give-back-past-pl
'The men gave those things back to one another.'

7.5. Cross-Over Constraint The cross-over constraint reaches across a sentence and down into the morphology of the verb. In this regard the argument component of the morphology is globally sensitive to the syntax, and might in fact be called syntactical morphology. The cross-over constraint explains why the absolutive relative index /yə-/ is so rarely used: it complicates things for the speaker. In (47) traces are noted by the customary *e*.

(47) Cross-over and zero

a. English form (from the jacket of an old Beatles album)

*He sang a song *which* I forget *how* it goes *e*.

b. Form with overt relative index in the absolutive

*ē₂ ē₁ ø₁-z-a-r-a₂-z-a-āa-ā+āa-a-ya-w sādaws^{hyt} t^haw₂ //
(how)₂ (song)₁ 3₁-what-dat-instr-dat₂-self-dat-mass-lie+active-
th v-past-comp how₂
yē₁-z-a-s^{hy}ə-z-ya-g^wə+pša-a-ya zə-warad-g^wara-r ø-ø-ə-ŋ^wa-a-ya
which₁-self-dat-deixis-I-cause-heart+forget-th v-past-adj
one-song-certain-abs 3-3-non pr-say-th v-past

*'He sang a song which I forget how it goes.'

c. Form with zero-relative index in the absolutive

ē₂ ē₁ ø₁-z-a-r-a₂-z-a-āa-ā+āa-a-ya-w sādaws^{hyt} t^haw₂ //
(how)₂ (song)₁ 3₁-what-dat-instr-dat₂-self-dat-mass-lie+active-
th v-past-comp how₂
ø₁-z-a-s^{hy}ə-z-ya-g^wə+pša-a-ya zə-warad-g^wara-r ø-ø-ə-ŋ^wa-a-ya
which₁-self-dat-deixis-I-cause-heart+forget-th v-past-adj one-
song-certain-abs 3-3-non pr-say-th v-past
'He sang a song, I forget how it goes.'

8. Conclusions and Proposals. A number of conclusions should be enunciated here. They offer important challenges for further linguistic

research into morphological theory.

8.1. Autolexical Theory. Autolexical syntax brings simplicity to such verbs if taken down into lexical structure. Verb morphology for languages such as Circassian is made up of interacting autonomous sub-levels. As noted at the outset there is no $G_a \times G_b$ interaction (multiple parsing) as one has between two autonomous levels, but merely a union of subcomponents. These sub-levels are: argument structure (syntactical morphology), derivational morphology (including compounding), clitic-like trains (inflectional morphology, including compounding), incorporation, and pragmatic morphology.

8.2. Governing principles The interaction of these sub-levels seems to be controlled by four principles:

First, what might be called information protocols rank morphemes by their content, so that time precedes manner precedes place.

Second, the position of head morphemes determines the direction in which an information protocol runs. Thus, the clitic train is right-headed because the syntax is right-headed, with place preceding manner preceding time, whereas the argument structure, being ergative, is left-headed, and thus the mirror image of syntax.

Third, the cross-over constraints are true of syntax and morphology, indeed of information encoding at all levels.

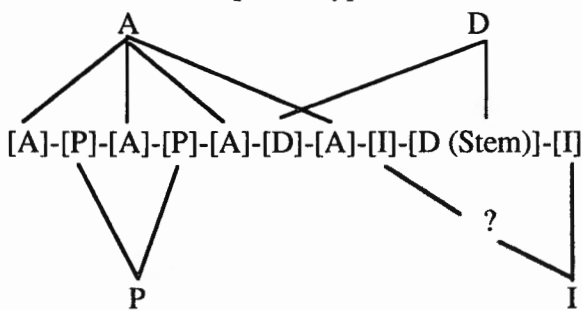
Fourth, scope constraints on logic and sense work on all levels as well. Thus, constraints become overall features of information encoding on all levels of grammar. This is predicted by Autolexical theory; which says that no one level can control it all.

The overall order is that in (48) (see below). Herein A = argument, P = pragmatical, D = derivational, and I = inflectional morphology.

The consequences for lexical phonology (morphology) (see Spencer 1991:105-119) are serious given the intercalation of levels in (48). In fact, in Colarusso 1992 I assigned adverbial argument indices to the pragmatic morphology merely to avoid the alternating argument—pragmatic morphology at the beginning of the verbal complex—but this was a Procrustean abuse of the data. Some form of this theory might be preserved, but this would necessitate suspending the bracketing erasure convention (Spencer 1991:113) and reinstituting some form of cyclic rule application to account for the alternation of argument and pragmatical morphology at the beginning of the verb complex. Such a theory would be a very weak form of lexical morphology, and yet

with the small exception of a few stem-initial inflections (44d, e) lexical morphology works well for the inflectional clitic train. Further, it works well for derivational morphology with the sole exception of the derivational preverbs. One might note similar variation in English between separable and non-separable intransitive prepositions that seem to have a derivational status (49a, b), and separable prepositions [\pm transitive], (49c, d). Therefore such separability of derivational morphology is not an exotic problem.

(48) Overall order of morpheme types



(49) English prepositions [put-on]

- a. Inseparable, derivational, [- transitive] preposition
 - (i) He is putting on airs.
 - (ii) *He is putting airs on
- b. Separable, derivational, [- transitive] preposition
 - (i) He is putting on a show.
 - (ii) He is putting a show on.
- c. Separable, syntactic, [- transitive] preposition
 - (i) He is putting on his coat
 - (ii) He is putting his coat on.
- d. Separable, syntactic, [+transitive] preposition

He put his coat on the table.

Clearly an interacting set of morphological (and phonological) sub-components with links to syntax (see Baker 1988b) and the lexicon, as well as to pragmatic, logical, and scope dimensions of the grammar is necessary to give a principled and complete account of morphology.

Autolexical theory, with its system of parallel processing, if modified to allow the union of subcomponents working within Cartesian products of full-sized components, offers the simplest and yet most powerful overall framework for building such a theory.

8.3. Zero Morphemes as Defaults. The results in (47) also strongly suggest that zero morphemes are not real (have no theoretical status) and that such forms might be best seen as default values of forms in a system of interpretation. If morphological zeros may not be real, then the same negative conclusion may be held for syntactic traces as well (for traces see Haegeman 1991:393-436). This would require a radical rethinking of current syntactic theories, somewhat along lines suggested by Mithun (1986) or, for syntax, by Cooper (1982).

8.4. Word Phrases. The study of the Circassian verb makes it useful to coin a new notion for morphology, that of the word phrase. Such an element shows more flexibility and expressive power than the usual word, but at the same times functions as a phonological, and presumably, cognitive unit. I would conjecture that only verbs or inflectional elements associated with verbs can be found as such word phrases. This fact would reflect the basic character of a verb in defining sentences other than copular ones.

8.5. Morphological Features. The complex verbs can show eight features that make them similar to whole sentences:

- (1) ordering/scope variation,
- (2) variation in roles for morphemes,
- (3) internal (hierarchical or linear) structure within morphemes,
- (4) homophony of morphemes,
- (5) transformations (or some topologically equivalent mechanism),
- (6) context sensitivity,
- (7) referential function (variables and generalised quantifiers), and
- (8) constraints similar to those operating on syntax.

8.6. Global Features of the Verb. The verbs are phonologically and cognitively a unit. They have an expressive capacity rivalling a sentence (lacking only the full referential power of a sentence). They have hierarchical or linear structure within their morphemes. They have a modicum of syntax-like flexibility, namely, varying morpheme-order. They are best described by a G_2 PS grammar with a dash of G_1 and G_0 (Wall 1972).⁶ They have nevertheless extreme productivity.

⁶I should add the proviso that unlike syntactic grammars, word grammars do

They can even surpass the sentence in so far as they can contain pragmatic material not necessarily found in the sentence. Therefore, such verbs are better termed "word phrases" or WPs. In a most general sense we may therefore look upon morphology as biologically distinct from syntax, but not formally distinct from it (both are G_0 's).

8.7. Effects of Complex Verbs. Such verbs are virtual microcosms of the sentence. They enrich the expressive capabilities of the language in that they alter discourse structure. Typically a discourse proceeds by laying out the referents by full sentences with nouns. Once the referential world is delineated, one then proceeds largely by verbs. As the interrelationships within a discourse grow more complex, so do these verbs, until they reach a rhetorical climax which is usually coincident with that of the narrative. These verbs provide an enormously efficient (rapid) means of delivering information. They therefore offer a degree of cognitive force and facility lacking in less endowed languages. As a young speaker of Abadzakh West Circassian, Mr. Adnan Saygili (Circassian name Zhazhy Adnan), once told me, "Circassian is a beautiful language. You can say everything in one word."

APPENDICES

1. Symbols and Abbreviations

| | | |
|-------|---|---------------------------------------|
| ǎ | = | vowel with primary stress |
| ǎ̂ | = | vowel with secondary stress |
| 3 | = | 'he,' 'she,' 'it,' 'him,' 'her,' 'it' |
| A | = | adverb node |
| abs | = | absolutive case |
| act | = | active |
| adv | = | adverb |
| aff | = | affirmative mood |
| afx | = | affix |
| C | = | complementizer |
| cause | = | causative |
| conj | = | conjunction |
| con | = | connective |
| dat | = | dative |

not seem to show extraction processes, unless that of Bella Coola affixation jumping (note 4, example D) might be usefully interpreted in this way, though a simple theory of clitics would seem to obviate this possibility.

| | | |
|--------|---|--|
| def | = | definite time |
| dir | = | direction |
| dist | = | distributed through space or time |
| dyn | = | dynamic |
| emph | = | emphatic |
| fut | = | future |
| ger | = | gerund (a participle-like form that can take case endings) |
| hor | = | horizon of interest |
| incept | = | inceptive (or change of state) |
| imp | = | imperative |
| inf | = | infinitive |
| INFL | = | inflection node (much like the old Aux(iliary) node) |
| intr | = | intransitive |
| io | = | indirect object |
| irreal | = | irrealis |
| loc | = | locative |
| m | = | marked |
| N | = | noun node |
| neg | = | negative |
| non pr | = | non-present tense vowel (/e/) |
| num | = | number (plural) |
| obl | = | oblique case |
| opt | = | optative |
| perf | = | perfect |
| pl | = | plural |
| poss | = | possessive |
| pot | = | potential |
| pred | = | predicative case |
| pres | = | present tense |
| pro | = | pronominal suffix |
| prog | = | progressive aspect |
| Q | = | question complementizer |
| recip | = | reciprocal |
| ref | = | referential |
| S | = | sentence node |
| subj | = | subjunctive |
| surf | = | surface |
| V | = | verb node |

2.Sound Systems

(A) Bzhedukh West Circassian segmental inventory

| | | | | | | | |
|-------------------|----------------|---|---|----|---|---|---|
| labial | p ^h | p | b | p' | f | m | w |
| dental (alveolar) | t ^h | t | d | t' | | n | |

| | | | | | | | |
|---------------------------|-----------------|----------------|----------------|-----------------|-------------------|----------------|-----------------|
| alveolar | c ^h | c | ʒ | c' | s | z | r |
| lateral | | | | | ʃ | ʒ | ʃ' |
| alveolo-palatal | | | | | š | ž | š' |
| rounded | č ^{hw} | č ^w | ʒ ^w | - | š ^w | ž ^w | š' ^w |
| palato-alveolar (laminal) | č ^{hy} | č ^y | ʒ ^y | č ^{'y} | š ^{hy} | ž ^y | |
| retroflexed | č ^h | č | ʒ | č' | š ^h | ž | |
| velar (palatal) | - | k | - | k' | (x ^h) | ḡ | y |
| rounded | k ^{hw} | k ^w | g ^w | k' ^w | x ^w | - | |
| uvular | q ^h | q | - | - | χ | ʁ | |
| rounded | q ^{hw} | q ^w | - | - | χ ^w | ʁ ^w | |
| pharyngeal (epiglottal) | | | | | ħ | ʕ | |
| laryngeal | | | | ʔ | h | | |
| rounded | | | | ʔ ^w | | | |

vowels ə a

(B) Kabardian segmental inventory

| | | | | | | | | |
|---------------------------|----------------|----------------|-----------------|-----------------|----------------|----|---|---|
| labial | p | b | p' | f | v | f' | m | w |
| dental (alveolar) | t | d | t' | | | | n | |
| alveolar | c | ʒ | c' | s | z | | | r |
| lateral | | | | ʃ | ʒ | ʃ' | | |
| alveolo-palatal | | | | š | ž | š' | | |
| palato-alveolar (laminal) | | | | š | ž | | | |
| palatal | k ^y | g ^y | k' ^y | x̂ | ḡ | | | y |
| rounded | k ^w | g ^w | k' ^w | x̂ ^w | - | | | |
| uvular | q | - | q' | χ | ʁ | | | |
| rounded | q ^w | - | q' ^w | χ ^w | ʁ ^w | | | |
| pharyngeal (epiglottal) | | | | ħ | ʕ | | | |
| laryngeal | | | ʔ | h | | | | |
| rounded | | | ʔ ^w | | | | | |

vowels ə a

REFERENCES

- Baker, Mark. 1988a. *Incorporation, a Theory of Grammatical Function Changing*. Chicago: University of Chicago.
- _____. 1988b. Morphology and Syntax. In *Morphology and Modularity: in Honour of Henk Schultink*. eds. Martin Everaert, Arnold Evers, Riny Huybregts, and Mieke Trommelen. 9-32. Dordrecht, Holland: Foris.
- Catford, J. C. 1984. Instrumental Data and Linguistic Phonetics. In *Topics in*

- Linguistic Phonetics in Honour of E. T. Uldall*, ed. Jo-Ann W. Higgs and Robin Thelwall. *Occasional Papers in Linguistics and Language Learning*, no. 9, March 1984. Ulster: The New University of Ulster. Pp. 23-48.
- Colarusso, John. 1992. *A Grammar of the Kabardian Language*. Calgary, Alberta: University of Calgary Press.
- _____. 1984. "Circassian /-qa-/ Meets Semantic Algebra and Ergativity." Aronson, Howard I. (ed.) *Folia Slavica, Papers from the Third Conference on the Non-Slavic Languages of the USSR* 7: 49-90.
- Cooper, Robin. 1982. "Binding in Whole-Wheat* Syntax (*unenriched with inaudibilia)." Jacobson, Pauline and Pullum, Geoffrey K. (eds.) *The Nature of Syntactic Representations*. Dordrecht, Boston, London: D. Reidel Publishing Co. Pp. 59-77.
- Davis, Philip W. and Saunders, Ross. 1980. *Bella Coola Texts, Heritage Record No. 10, British Columbia Provincial Museum*. Victoria, B.C.: British Columbia Provincial Museum.
- Dumézil, Georges and Namitok, Aytek. 1939. *Fables de Tsey Ibrahim (tcherkesse occidental)*. Paris: Paul Geuthner.
- Hadaghat'la, Asker [ħadayaħ'a askar. 1968. Nartxer [The Narts], vol. 1. Majkop: Adyghe Naučna Issledovatel'ska Institut, Sektor Literatura. (in Circassian).
- Haegeman, Liliane. 1991. *Introduction to Government and Binding Theory*. London and Cambridge, Massachusetts: Basil Blackwell.
- Kardanov, B. M. 1955. "Kratkij grammatičeskij očerk kabardino-čerkesskogo jazyka." B. M. Kardanov and A. T. Bičoev (eds.), *Russko-kabardinsko čerkesskij slovar'*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej. Pp. 993-1054.
- Keraševa, Z. I. 1960. "Kratkij grammatičeskij očerk adygjskogo jazyka." X. D. Vodoždokov (ed.), *Russko-adygjskij slovar'*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej. Pp. 1059-1098.
- McCawley, James D. 1988. *The Syntactic Phenomena of English*. Chicago: University of Chicago.
- Mithun, Marianne. 1986. When Zero Isn't There. Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj, and Deborah Feder (eds.) *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. Pp. 195-211.
- Paris, Catherine. 1974. *La princesse Kahraman, Contes d'Anatolie en dialecte chapsough (tcherkesse occidental)*. Paris: Société d'Études Linguistique et Anthropologiques de France (SELAF).
- Sadock, Jerrold M. 1991. *Autolexical Syntax*. Chicago: University of Chicago Press.
- Scancarelli, Janine. 1986. Pragmatic Roles in Cherokee Grammar. Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj, and Deborah Feder (eds.) *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. Pp. 225-234.

- Sells, Peter. 1985. *Lectures on Contemporary Syntactic Theories*. Stanford, California: Stanford University, Center for the Study of Language and Information.
- Smessaert, Hans. 1988. An Autolexical Approach to Pronominal Cliticization in West Flemish. Master's Thesis, University of Chicago.
- Smeets, Riëks. 1983. "On Location and Direction in Circassian: Five Directional Suffixes," *Folia Slavica* 5: 384-394.
- _____. 1984. *Studies in West Circassian Phonology and Morphology*. Leiden: Hakuchi Press.
- Spencer, Andrew. 1991. *Morphological Theory*. London and Cambridge, Massachusetts: Basil Blackwell.
- Wall, Robert. 1972. *Introduction to Mathematical Linguistics*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

AN EVALUATION OF CHUKCHEE ORTHOGRAPHY

Bernard Comrie

University of Southern California

In this paper, my aim is to carry out an evaluation of the current orthography of Chukchee, in particular assessing the extent to which this orthography provides a good fit to the structure of the language versus the extent to which it is influenced by Russian orthographic conventions. Chukchee is the major language of the Chukotko-Kamchatkan family, and is spoken by some 11,000 people in the Chukotka peninsula in northeastern Siberia. The current Cyrillic orthography was introduced in 1936, as part of the general Soviet policy of introducing Cyrillic orthographies for the languages of the USSR. Previously, from around 1931, a Latin orthography had been used. For the data on Chukchee phonemics and orthography, see Skorik (1961).

In principle, an orthographic system can serve (at least) two purposes. The first is correspondence to the structure of the language in question. The second is a cultural function. I will elucidate and exemplify these functions before turning specifically to the evaluation of the Chukchee orthographic system.

The most neutral sense in which 'correspondence to the structure of the language' is usually taken, in speaking of an alphabetic orthography, is that the orthography should represent all and only the taxonomic phonemes of the language, with if possible a one-one correspondence between phonemes and orthographic symbols (though in principle digraphs might be permitted if their use does not lead to ambiguity). Orthographies recently devised or currently being devised for many languages of Africa and the Pacific area illustrate this possibility, for instance Swahili (with some digraphs, e.g., *ch*). Of course, an orthography might also be required to take into account other levels of analysis, in addition to or instead of the taxonomic phonemic level, for instance, morphophonemics. Thus, one principle of Russian orthography is that vowels in a given lexeme should always receive the same representation, even though their phonemic value may change as a result of shifting stress, e.g., голова /gala'va/ 'head,' nominative plural головы /'golavi/, genitive plural голов /ga'lof/, where the orthographic vowel *o* is established consistently on the basis of morphophonemic identity (and likewise the stem-final *в* (= *v*), which

is devoiced in the genitive plural). With respect to the current Chukchee orthography, however, to the extent that it provides a good fit to the structure of the language, this fit is to the taxonomic phonemic representation. Thus, where alternations take place (and such are frequent in Chukchee), the phonemic result rather than the morphophonemic identity is indicated in the orthography, e.g., йып-ык /jəpək/ 'to put on,' йым-нэн /jəmnen/ 'he put it on,' with automatic partial assimilation of /p/ before /n/.

In speaking of the cultural functions of an orthography, there are several factors that could in principle come into play, but for present purposes only one is relevant: an orthography can, by assimilating to or distancing itself from the orthography of some other language (or earlier stage of the same language) provide a means of identifying with or distancing itself from the culture associated with that other language or earlier stage. Thus English orthography, which provides a rather poor fit to the structure of any given contemporary dialect, enables identification to be made both across dialects and with earlier stages of the language. The use to which the Latin letter *j* is put in many recently devised orthographies provides another interesting illustration: in areas with English cultural influence, *j* usually represents an affricate or stop in the palatal region; in areas with Spanish cultural influence, it is more likely to represent a voiceless velar fricative; in areas with Dutch cultural influence, it is more likely to represent a palatal semivowel (as in the pre-reform orthography of Indonesian). Serbo-Croatian can be written either in the Cyrillic script or in the Latin script, thus providing continuities to different cultural traditions, and also enabling an internal cultural differentiation to be made within the one language. The examples of Latin *j* and Serbo-Croatian show, incidentally, that it is not necessary for cultural and structural factors to be in conflict: in Swahili there is one-one correspondence between the letter *j* and the phoneme /j/; Serbo-Croatian, whether written in the Cyrillic or Latin script, shows a close fit between orthography and phonemics.

In the case of Chukchee, the only cultural factor which will be relevant is the extent of assimilation of Chukchee orthography to orthographic practices of Russian. Thus, in what follows, I will examine the extent to which the orthography provides a fit to the structure of the Chukchee language versus the extent to which this fit is disrupted

by adherence to practices taken over from Russian. Given the social realities of life in the Russian Federation, there is of course a case to be made for admitting at least some similarities between Chukchee and Russian orthographies. Children in Chukchee schools are also taught Russian, which is the sole medium of instruction after the first few grades, so that there is some facilitation of the educational process if the same orthographic devices can be used for both languages, with the result that children need assimilate only one rather than two sets of principles. This is one motivation for the use of the Cyrillic script. An analogy would be the practice in the orthographies of some languages spoken in Latin American countries where the dominant language is Spanish, of using the Spanish convention for representing velars, where for instance /k/ is written *c* before back vowels but *qu* before front vowels. While this does to some extent disrupt the fit between orthography and structure, the disruption is usually quite small. If, however, the disruption is greater, then the value of an orthography as a correspondent to the structure of the language can be seriously undermined.

TABLE 1
CHUKCHEE PHONEMES AND ORTHOGRAPHY

| IPA | Latin transcription | Cyrillic orthography |
|-----|---------------------|----------------------|
| i | i | и |
| u | u | у (ю) |
| e | e | э (е) |
| o | o | о (ё) |
| a | a | а (я) |
| ə | ə | ы |
| p | p | п |
| t | t | т |
| k | k | к |
| q | q | к |
| ʔ | ʔ | ь (ь, ') |
| tʃ | c | ч (ç) |
| β | v | в |
| ɣ | g | г |
| ɬ | l | л |

| | | |
|---|---|----------------------|
| ɬ | ɾ | ɖ |
| j | j | й (ь, ь, е, ё, ю, я) |
| m | m | м |
| n | n | н |
| ŋ | ŋ | ң |

Table 1 gives the phonemes of Chukchee in the International Phonetic Association transcription (first column), followed by the Latin transcription to be used in the remainder of this article (which simply replaces typographically complex symbols by simpler ones where possible). The third column gives the Cyrillic symbol corresponding to each phoneme. The main entry in the third column gives the Cyrillic symbol which most usually corresponds to that phoneme; in parentheses are given other Cyrillic symbols which, under certain circumstances (detailed below), represent, or participate in the representation of, that phoneme. If one looks simply at the basic representations, then the Cyrillic and Latin transcriptions are essentially equivalent: both have one-one correspondence between phoneme and letter, both avoid where possible the use of typographically unusual symbols where a given Chukchee phoneme is phonetically similar to and does not contrast with the value of a typographically simpler symbol. (The only special symbols created for the Chukchee alphabet are к (also written к') and ң (also written н'); in Russian, the following letters have slightly different values: ы /ɨ/, ч /tʃ/, в /v/, г /g/, л /l/, р /r/; for ь, ь see below.)

Let us now turn to the use of the Cyrillic symbols given in parentheses in Table 1. The discussion will be restricted to native and fully assimilated words; loans from Russian are simply written as they would be in Russian, an unequivocal intrusion of Russian orthographic practices.

In Chukchee, there is overall relatively little positional allophony, but one of the striking exceptions concerns the phoneme /c/, which is realized as a fricative, [s], before /q/, e.g., [resqən] 'roof.' The closest Russian phonetic equivalent to this fricative allophone would be the voiceless alveolar fricative /s/, for which the Cyrillic symbol is c. In Chukchee orthography, in this one case alone an allophonic variant is represented explicitly, i.e., the phoneme /c/ is written c before /q/, as in рэскын. Since the orthographic symbols otherwise correspond essentially one-one to phonemes, the only motivation for this

differentiation is to assimilate Chukchee orthography to Russian, where the affricate /tʃ/ and the fricative /s/ are distinct phonemes.

Russian has a phonemic opposition between palatalized and nonpalatalized consonants, and many of the complexities of Russian orthography are concerned with the representation of this phonemic distinction without having to devise a separate letter for each palatalized consonant. In particular, where a consonant is followed by a vowel in Russian, the palatalization value of the consonant is shown by the following vowel symbol, as follows: та /ta/, тэ /te/, то /to/, ту /tu/, but тя /tʲa/, те /tʲe/, тө /tʲo/, тю /tʲu/. (In Russian, a case can be made for treating the distinction between ты and ти likewise, i.e., analyzing ты as /ti/ and ти as /ʲi/. However, there are also reasons for analyzing ы and и as representing different phonemes, respectively /ɨ/ and /i/. Whatever the best analysis for Russian, in Chukchee ы and и clearly represent different phonemes, thus corresponding to the second analysis for Russian proposed above, and indeed leading to sequences of letters impossible in Russian, e.g., чы /cə/ in чавчыв /cavcəv/ 'reindeer-breeding Chukchee.') Chukchee has no phonemic opposition of palatalization. However, two of its consonants are phonetically palatalized, namely /c/ and /l/. In Russian, there is a phonemic opposition between nonpalatalized /t/ and palatalized /tʲ/, and thus the closest Russian phonetic correspondent to Chukchee /l/ is in fact the palatalized /lʲ/. Because of this phonetic correspondence, the principle was introduced into Chukchee orthography that when /l/ is followed by a vowel, the orthographic representation should correspond to that for a palatalized consonant in Russian, i.e. ля, ле, лё, лю, e.g., лялёлгын /laləlgən/, plural лелют /ləlut/ 'beard.' (In Russian, palatalization before a consonant or pause is indicated by appending the letter ь after the consonant, e.g., ль for /lʲ/; this orthographic device was not carried over to Chukchee /l/ in syllable-final position, which is simply written л, e.g., ейвэл /jejvel/ 'orphan'; presumably because ь after л has another function, for which see below.)

Russian /tʃ/ is always palatalized, and as with the other consonants which do not participate in the palatalization opposition, vowels after ч are represented as follows: ча, че, чо (in certain forms also чё), чу. These representations were simply taken straight into Chukchee, e.g., чавчыв /cavcəv/ 'reindeer-breeding Chukchee,' чейвэ /cejve/ 'on foot.' The different representation of vowels after different consonants

corresponds to nothing in the structure of Chukchee.

One of the other major complications of Russian orthography is the representation of the phoneme /j/. The symbol й is used only preconsonantly and wordfinally. Between vowels or word-initially before a vowel, the representation of /j/ is combined with that of the following vowel, so that the letters ('yotated' vowels) in this environment have the following values: я /ja/, е /je/, ё /jo/, ю /ju/. After a consonant, /j/ is represented in a particularly complex way: after the consonant follows one of the two letters ь or ъ (the latter is used after certain prefixes only, which in their citation form end in a nonpalatalized consonant), then the vowel symbol from the yotated series mentioned above, i.e., we find representations like тья or тъя for /tja/. The sequence /tji/ is represented by тьи. Essentially these same practices, with the exception of the last-named, are carried over into Chukchee.

Chukchee /j/ is represented as й syllable-finally, e.g., аймын /ajmən/ 'hole in ice.' It is also represented as й before /ə/ and /i/, e.g., йымэк /jəmek/ 'to hang (tr.),' йитык /jitək/ 'to drip' (in Russian, the sequence /ji/ is impossible, while a phonemically distinct /ji/ occurs only after consonants, as discussed in the previous paragraph). Otherwise, between vowels and word-initially, Chukchee /j/ is represented as in Russian by using the letters я, е, ё, ю, e.g., янот /janot/ 'earlier,' юнрык /junrək/ 'to select.'

For /j/ after a consonant, Chukchee again uses basically the Russian system, though with ъ rather than ь except after /l/ and /c/--another concession to Russian palatalization--e.g., памъят /pamjat/ 'woolen stockings,' маляа /maljaa/ 'rather distant.' Once again, the representation of /j/ before /i/ and /ə/ does not follow this system, but rather provides a one-one fit to the phoneme sequence, e.g., эвйикэ /evjike/ 'silently' (not, as one might expect from Russian orthographic practice, *эвыикэ), йыныйн /jənɲən/ 'fire.'

Although the Russian system is complex, it has an internal logic and consistency as applied to Russian. When applied to Chukchee, however, not only does it disrupt the fit between phoneme and letter, but it also provides a system lacking in internal consistency: the fact that Russian does provide models for representing /ja/, /je/, /jo/, /ju/, but does not provide comparable models for /ji/ and /jə/, means that Chukchee is left with an inconsistent compromise between the Russian system (where

Russian provides a model) and a phonemic orthography (where Russian does not provide a model). In fact, the situation becomes even worse if we consider one of the major morphophonemic alternations in Chukchee, vowel harmony. The Chukchee vowels divide into two sets, called dominant and recessive, as in Table 2 (note that /e/ is both the dominant counterpart of /i/ and the recessive counterpart of /a/, while /ə/ can be dominant or recessive).

TABLE 2
CHUKCHEE VOWEL HARMONY

| | | | | |
|-----------|---|---|---|---|
| Dominant | e | o | a | ə |
| Recessive | i | u | е | ə |

Chukchee words frequently consist of several morphemes, and if in a given word at least one morpheme contains dominant vowels, then all recessive vowels within that word must be replaced by their dominant counterparts, e.g., /nelg-ən/ '(animal) skin' but /ga-nalg-əma/ '(animal) skin (Comitative).' In general, the effects of vowel harmony can be seen readily in the orthography through the alternation of the letters и with э, у with о, э with а, e.g., нэлгын, ганалгыма. After /l/, the orthographic alternations are slightly different, but still transparent (и with е, ю with ё, е with я), and likewise after /c/ (и with е, у with о, е with а). After /j/, the alternations are again transparent for the alternations of ю with ё and е with я, but become cumbersome with the /i/ - /e/ alternation, given that Chukchee uses йи for /ji/ but е for /je/. The orthographic shape of words containing the /i/ - /e/ alternation after /j/ is thus grossly distorted by the effect of vowel harmony, e.g., йигйит /jigjit/ 'small intestines,' singular егъелгын /jegjelgən/, with the harmonically dominant suffix /-lgən/.

Finally, let us consider the representation of /ʔ/. Russian lacks a glottal stop, so the assignment of the letter ь to this value is in a sense arbitrary, though an economical use of existing typographical symbols. The symbol ь is used in Russian, as noted above, in the representation of /j/ after nonpalatalized consonants. This means that the distribution of ь in Russian is quite restricted: it only occurs after nonpalatalized consonants, and moreover only before the vowel symbols я, е, ё, and ю. The second of these restrictions was not carried over into Chukchee

(it would not have been possible to do so, since the sequence /Cʔj/ is nonphonotactic in Chukchee). However, the first restriction was carried over. Thus we find a close fit between orthography and phonemics for /ʔ/ only where the glottal stop follows a (phonetically) nonpalatalized consonant, e.g., ирѡын /irʔən/ 'topcoat.' After palatalized consonants, the counterpart of ѡ in Russian is ѣ; this correspondence is carried over into Chukchee, so that after /c/ and /l/ one writes ѣ rather than ѡ, e.g., ымыльѡ /əməlʔo/ 'all.'

In Chukchee, the glottal stop can also occur without a preceding consonant, either word-initially, e.g., /ʔaacek/ 'youth,' /ʔigən/ 'wolf,' or intervocalically, e.g., /ʔiʔin/ 'sky,' or syllable-finally, e.g., /reʔtək/ 'what did you-all do?,' /iʔrek/ 'to cross (a river).' In these environments, it was apparently felt too un-Russian to use either ѡ or ѣ. For the sequence /CVʔ/, the expedient selected is to write ѡ before, rather than after, the vowel, e.g., рѡэтык, although this does introduce an ambiguity (note that none of the previously discussed complications in Chukchee orthography prevents going unequivocally from phonemic representation to orthography and vice versa), since this could in principle also represent /rʔetək/, cf. рѡэт /rʔet/ 'way.' In the other cases, an apostrophe is used, invariably placed after the vowel, even when phonetically it precedes, e.g., а'ачек, и'гын, и'и'н, и'рэк; again, this introduces a potential ambiguity, since V' can represent either /ʔV/ or /Vʔ/. (In fairness, it should be noted that some dialects of Chukchee apparently do not distinguish phonetically the position of a glottal stop relative to an adjacent vowel or vowels, having rather glottalization of the vowel(s), although these are not the dialects on which the standard language is supposed to be based.) The representation of /ʔ/ in Chukchee is thus extremely complex, not because of any inherent properties of Chukchee, but simply because of graphotactic restrictions transferred from Russian; this case seems particularly unnecessary given that the symbol ѡ is used in Chukchee for a value completely different from that which it has in Russian.

Since our discussion has been primarily in terms going from sound to writing, Table 3 summarizes the information in the reverse direction, listing the letters of the Chukchee alphabet (in alphabetical order, excluding those used only in unassimilated loans from Russian), plus the apostrophe, with their usual phonemic equivalent and secondary phonemic equivalents (including instances where the letter is used as

part of the representation of a given phoneme) in parentheses.

TABLE 3
CHUKCHEE ALPHABET WITH CORRESPONDING PHONEMES

| Cyrillic | Phoneme |
|----------|---------|
| а | a |
| в | v |
| г | g |
| е | je (e) |
| ё | jo (o) |
| и | i |
| й | j |
| к | k |
| қ | q |
| л | l |
| м | m |
| н | n |
| ң | ŋ |
| о | o |
| п | p |
| р | r |
| с | s |
| т | t |
| у | u |
| ч | ʃ |
| ъ | ʔ (j) |
| ы | ə |
| ь | (ʔ, j) |
| э | e |
| ю | ju (u) |
| я | ja (a) |
| , | ʔ |

Given that an orthography has both structural and cultural functions, there is nothing inherently wrong with transferring orthographic practices from one language to another with which it is in cultural contact. In the case of Chukchee, however, the question arises seriously

of whether the current Cyrillic orthography does not make too many concessions to Russian orthographic principles, thereby undermining its value as a carrier in written form of the structure of the Chukchee language.

REFERENCE

- Skorik, P.Ja. 1961. *Grammatika čukotskogo jazyka*. Part I. Moscow/Leningrad: Izdatel'stvo Akademii nauk SSSR.

RUSSIAN AND ROMANIAN INTERTWINED: THE LEGACY THAT IS MOLDAVIAN¹

Donald L. Dyer
University of Mississippi

"Two trends apparently will survive this decade and, most likely, the rest of the century. The first is the re-worked—yet still determined—Soviet separatist policy with regard to the Moldavian language, as evidenced by two large and recent treatises on the subject appearing in just the first few years of this decade: Il'jašenko 1983 and Korlètjanu 1983. As additional proof of continued Soviet separatist sentiment, I offer some statistics culled from Talmackaja 1978, a bibliography of Soviet linguistic works on Moldavian in print through the late 1970's. Talmackaja lists 214 entries on such works. ... A full 85% of the works (182) were published in the 1970's, an overwhelming majority being issued specifically in 1975 and 1976. No doubt the watchdog attitude toward Soviet language policy in the MSSR harbored by a growing number of Western scholars will likewise persist. This will be the other surviving trend." (Dyer 1990a:103)

1. Introduction. In 1987, the language situation in the Moldavian Soviet Socialist Republic (hereafter "MSSR") had an unclear future. Since that time, the Soviet Union has set upon a path of restructuring its society and, as a result of this, has experienced turbulent inner strife, particularly in its outer republics, where various ethnic groups have attempted to assert their own authority, politically establishing for themselves some sense of union-internal autonomy. In some cases, republics have declared outright independence from the Soviet Union (the Baltics, for example). In the MSSR, things have been difficult as well. The Moldavians, the predominant ethnic group in the region, through a series of political maneuverings in the second half of the 1980s, took control of the local government of the MSSR (Crowther, to appear; *Glasul Națiunii*, July 4, 1991). Their earlier calls for political independence from Moscow culminated in a number of deaths that

¹The text of this paper has changed little since its presentation at the Seventh International Conference on the Non-Slavic Languages of the USSR (May 22, 1991, Chicago, Illinois). I wish to thank Michael Bruchis of Tel-Aviv University, Professor William Crowther of the University of North Carolina at Greensboro, Professor Gary Toops of Wichita State University and Professor Kostas Kazazis of the University of Chicago for assistance provided and advice proffered during the production of this manuscript.

resulted from various ethnic and political clashes on the streets of the cities and in the villages (*Jackson Clarion-Ledger*, October 27, November 3, 17, December 17, 1990; *Memphis Commercial-Appeal*, October 28, 1990). So fragmented now are the people's allegiances in this region, and so confused have the issues of politics, ethnicity and language become, that the Gagauz, a minority of Turkic-speaking Orthodox Christians living in the MSSR, have declared for themselves a new republic—one that lies within the MSSR and whose provisional government is housed in Tiraspol (*New York Times*, September 15, 1990).

2. Flashpoint—The MSSR in the Summer of 1989. In August of 1989, hundreds of thousands of ethnic Moldavians rallied in Kishinev to demand the adoption of their language, Moldavian, as the official language of the MSSR. Soon thereafter, the Moldavian legislature voted 300-43 in favor of such adoption, and a law to the effect became operational on August 30, 1989. Additionally by decree, as of 1995 all official business in the republic would be conducted in Moldavian. A return to the Latin alphabet, last used between the two world wars, was effective immediately (*Chicago Tribune*, August 28, 30, 1989; *New York Times*, September 15, October 17, 1990).

Spearheading the move to have Moldavian become the official language of the MSSR was the Peoples' Front, a group of Moldavian independence-seeking separatists. The political and ethnic suggestions of this movement were not lost on another organization, *Edinstvo*. The latter group, an anti-separatist group of Russians, Ukrainians and Gagauz, had counterdemonstrated against the massing Moldavians in August of 1989 (Crowther, to appear; *Chicago Tribune*, August 28, 1989). *Edinstvo* continues to oppose any kind of separatist movement in Moldavia, as well as the official language status of Moldavian. Their rather formidable political might and their extensive breadth of support were obvious during the large-scale factory shutdowns of the summer of 1989, shutdowns that took place in Moldavian cities with large Russian- or Ukrainian-speaking populations, and which crippled the MSSR's economy for weeks (*Chicago Tribune*, August 30, 1989).

In early 1990, shifts in political sentiment and ethnic polarization, as well as the broader, sweeping changes taking place in the Soviet Union itself, brought the Peoples' Front to power in the legislature of the MSSR. A mix of Moldavian reformers (including hard-line Romanian

nationalists) and moderates who enlisted the help of Communist party faithfuls forged a strong alliance, an event that was quickly deemed unsatisfactory by minorities in the region (Russians, Gagauz and others). The Peoples' Front remains in power today, despite attempts by Romanian nationalists to derail the reform process altogether by calling for outright secession of Moldavia from the Soviet Union for the purpose of reunification with Romania (Crowther, to appear).

On April 24, 1990, Mikhail Gorbachev, recognizing that the language issue in Moldavia was becoming more and more dangerous, declared Russian the official and international language of the Soviet Union in an attempt to soften growing feelings of dissent. In Michael Bruchis' words, although "[i]n their struggle for the status of Moldavian and the Latin alphabet the ... [Moldavians] have indeed gained an impressive victory[, t]heir victory ... was reduced to zero by Gorbachev's decree ..."²

What appeared certain in the summer of 1991 is that although supporters of an official Moldavian have been successful on one level, under current policy, they will not legally be allowed to demand that non-Moldavian speakers in the MSSR function in official capacities in Moldavian. With matters still unresolved, Russians living in the MSSR continue to fight with Moldavians, Gagauz with both Russians and Moldavians, and Ukrainians with all the former. The language situation now is particularly complicated, but in 1991 in a way very different than it was in 1987.

3. Old Issues, Old Problems. Before perestroika and glasnost, linguists faced one basic question: "Why does language policy in the Soviet Union continue to support the line that Moldavian is a separate Romance language, the only member of the eastern group of the eastern branch of Romance (*New York Times*, February 25, 1989)?" By the early 1980s, we had seen a long succession of linguistic works, beginning with Magdan's *Gramatica moldovenească* in 1929, including Sergievskij's 1939 treatise *Moldavskie ètjudy*, and culminating in the Academy grammar of Moldavian (volume one by N.G. Korlètjanu, 1969 and volume two by A.M. Dyrul and A.I. Čobanu, 1970) that decried the linguistic autonomy of the Moldavian language and its distinctness from Romanian. The only proof of its distinctness was

²Personal communication, June 16, 1990.

offered in the form of numerous lexical differences between Moldavian and Romanian, despite the fact that as I show in Dyer (to appear), there exist a number of clearly discernible phonological and morphological differences between Moldavian (viz., Moldavian Daco-Romanian) and Muntenian Daco-Romanian, the basis for standard Romanian, that could have been utilized more productively to distinguish the two dialects. In fact, if the Kishinev dialectal features of Moldavian—those found in the heart of the Moldavian Daco-Romanian speech region—had been used as the base for “literary” Moldavian, we would have seen, at least on paper, two distinct speech varieties.³

Literary Moldavian, as presented in Korlètjanu 1969 and Dyrul and Čobanu 1970, is a “paper lion,” a virtual copy of literary Romanian, as it is presented in *Gramatica limbii române, I, Fonetica și morfologia*, 1963, the Academy grammar. Recent linguistic and politico-linguistic analyses of, and various works on, Moldavian quite clearly have shown that literary Moldavian, as packaged by the Soviets, was no more than literary Romanian written in Cyrillic (Bruchis 1979, 1980, 1982, 1983, 1984a, 1984b, 1987, 1988; Kazazis 1982; Deletant 1989; Dyer 1989, 1990a and to appear; Crowther, to appear). In fact, even the Communist Party of the MSSR admitted, in January of 1989, that Moldavian and Romanian are “identical” languages (*New York Times*, February 25, 1989).

4. New Issues, New Problems.

4.1. Moldavian as an Official Language. The issue at hand these days is no longer whether Moldavian is Romanian. Those who sought recognition of the truth that Soviet language policy was attempting to isolate the Moldavians from their kindred, the Romanians, through the use of geographic borders and the manipulation of their alphabets are no longer concerned with this issue. A series of events which began in the summer of 1989 (Dyer 1990b), which included demonstrations by Moldavians, Russians, Ukrainians, and Gagauz, as well as decrees by both Moscow and the Moldavian Congress in Kishinev, have produced a deep, at times violent, ethnic division of peoples. Among the results of such internal trouble has been Russian’s declaration as the official

³One need only consult the two dialectal atlases for this regional speech continuum to see how true this is: for Moldavia; specifically, Udler 1968, and for Romania-proper, Weigand 1909.

language of the Soviet Union, an action that complicates even further the language situation in the MSSR.

Moldavian independentists contend that Moldavian rightfully should be declared the official language of the MSSR. These people and their language do in fact stand as a dominating presence in the republic, despite the ever-present—and growing—Russian-speaking population. Though Moldavian independentists no longer feel threatened by the “Moldavian-is-not-Romanian” issue, they are now deeply concerned with a “Moldavian-should-not-be-displaced-by-Russian” issue. If Moldavian were to be declared the republic’s official language, despite denials in public about their actual intentions, Moldavian supporters eventually would be able to use the official language as a wedge to be driven between themselves and Moscow.

4.2. The Russian Influence on Moldavian. The concern that the status of the Russian language in the MSSR will grow is legitimate. The intent of Soviet language policy to divide the Romanian speakers of this general geographic region was a reality for half a century. The Moldavian dialect of Daco-Romanian was under attack from Russian for some 60 years, and the official line—though ignored by western, even some Soviet linguists—was that Moldavian, a language which after all was written in Cyrillic, was not Romanian. Studies as early as Tagliavani 1958, however, and densely throughout the 1980s, exposed this claim, upon linguistic analysis, as nothing more than political game-playing. On the front lines though, this particular brand of Romanian *was* being changed.

Despite claims by Soviet linguists until as recently as the mid-1980s, Moldavian Daco-Romanian, the regional speech variety of Romanian found in the MSSR, has always been but one recognized dialect of standard Romanian. Kishinev Daco-Romanian (Moldavian) differs from Muntenian (literary) Daco-Romanian in twelve minor, yet distinguishable, phonological and morphological features (Dyer, to appear). However, the influence of Russian on Moldavian—largely a result of the bilingualism found in this area—has been felt. Phonologically, the soft Romanian fricatives [š, ž] have been hardened as a result of their contact with hard Russian [š, ž] (Kazazis 1982: 227).⁴ Numerous Russian calques, including stump compounds, have

⁴ In fact, it should be understood that this hardening has taken place not only in

developed in Moldavian (Dyer 1990a:101). The Moldavian lexicon has been flooded with Russian vocabulary items. Today, this remains the major difference between the two varieties of Romanian.⁵ Syntactic influence on Moldavian has been negligible (Dyer 1989).

4.3. "Moldovan." For a limited time after the demonstrations and decrees of 1989, major publications in the MSSR were printed in both Latin and Cyrillic script, but in 1991, major journals of every kind are printed not only in Moldavian in Latin script, but in *Romanian*. Take, for example, the widely distributed literary journals *Revistă de lingvistică și știință literară* and *Nistru. Revistă de lingvistică și știință literară* had been published since 1958 as Лимба ши литература молдовеняскэ. Volume 5 of the 1990 issues hails its readers in the foreward with *Să scriem corect cu alfabetul latin!* 'Let us write correctly in the Latin alphabet!' *Nistru*, published since 1931, in addition to changing its type, has changed its name—to *Basarabia*. The editors in issue 9 of *Basarabia* (1990) argue for the name change with text and verse, claiming that *Nistru* is after all a Ukrainian name (and, thus, undesirable); a name such as *Basarabia* is purely Moldavian.

Basarabia a existat, există și mereu va exista în ciuda tuturor vigilenților.
Moldova, să nu uităm, începe dincolo de Prut. 'Bessarabia was, is
and will always be[,] in spite of all those who would keep vigilance.
Moldova, let us not forget, begins beyond the Prut [River].'

A map from a June 28, 1940 edition of *Izvestija* is included in the introductory *Argument* section of this issue. The map shows the Northern Bucovina and Bessarabian portions of Romania that historically had been lost to the USSR.

Both *Revistă de lingvistică și știință literară* and *Nistru* carry the interesting statement *na rumînscom iazîke*, 'in Romanian.' Collectively, these and other journals attempt not only to clarify for the ethnic Moldavians what language they speak, but to remind them of to what country they once belonged.

the Moldavian of the MSSR, but also in certain areas of the Moldavian Daco-Romanian dialect region of Romania-proper (Kostas Kazazis, personal communication).

⁵Korlètjanu 1983, for example, gives us a full 28 pages of Moldavian-specific vocabulary not found in Daco-Romanian.

5. Historical Perspectives and the Future of Moldavian.

5.1. Moldavian under Siege. Decrees of 1989 sought, among other things, to declare "Moldovan" the republic's official language.⁶ Had this been effected, Russian and Ukrainian speakers eventually would have been forced to become bilingual, many at relatively late ages in their lives. Official jobs, in particular, would have had as a prerequisite for employment native language ability in Moldavian. This issue continues to be inflamed (*Argumenty i fakty*, April 21-27, July 14-20, August 18-24, 1990; *New York Times*, October 7, 1990).

The Romanians/Moldavians have always been a plurality in this region. As early as the end of the nineteenth century, they made up 47.6% of the population of Bessarabia; 19.6% of the population was Ukrainian, 11.8% Jewish, 8% Russian, 5.3% Bulgarian, 3.1% Polish and 2.9% Gagauz. By the inter-war period of the twentieth century, the demographics of this region had changed substantially. In Romanian Bessarabia, the Romanian/Moldavian and Russian populations had increased significantly—to 56.2% and 12.3%, respectively. In the Autonomous Soviet Socialist Republic of Moldavia on the left bank of the Dniestr River, Romanians/Moldavians were in the minority, making up just over 30% of the population. Ukrainians represented a plurality at 48.5% of the population (Crowther, to appear). With the formation of the Moldavian Soviet Socialist Republic in 1940 and the forced relocation of Moldavians to Romania, as well as non-Romanian immigration to the MSSR from other republics, by 1970, the Moldavian population in the MSSR was declining. In urban areas in particular their population had been reduced to 35%, while the Russian population had risen to 28.8%. Moldavians were being relegated to the countryside; Russian speakers were taking over the big cities. In 1959, for example, Romanians were in a majority in only one of the ten largest cities—Orhei—and this was with a mere majority of 50.7% (Crowther, to appear).

By 1979, the composite MSSR was still 63.9% ethnic Moldavian,

⁶What was once called "Moldavian," is in fact being referred to as "Moldovan." Moreover, Moldavia is now being called "Moldova." The new names give the appearance of being politically neutral. This no doubt explains their recent emergence and present popularity. Interestingly, the International Research and Exchanges Board this year began to support a *Moldavian* language program in Kishinev (Cynthia Graves, personal communication, February 4, 1991).

the majority of Moldavians inhabiting rural areas. The linguistic complexion of the republic during this century has been quite diversified. The MSSR offers up a variety of ethnic and imported languages, truly one of the most diverse of the Soviet Union's republics with regard to language and culture. Censuses taken over the last fifty years have noted the presence of the following: Moldavian (Romanian), Ukrainian, Russian, Gagauz, Yiddish, Bulgarian, Gypsy, Belorussian, Polish, Armenian, Tatar and "other." One census taken in 1959 actually lists both Moldavian and Romanian as ethnic languages found in the MSSR. Subsequent sources do not list "Romanian." Censuses taken in 1959, 1970, and 1979 indicate that of the people calling themselves Moldavians (again, a rather consistent 63-65% of the overall population), a solid 96-98% considered Moldavian their native (first) language. Only 1.3%, 2.0%, and 3.3%, respectively, in those years considered Russian their native tongue (Deletant 1989:203-204). Rural Moldavians were more faithful to the language than those in urban areas. In 1970, for example, only 0.5% of rural Moldavians called Russian their native language, while 9.4% of urban Moldavians considered it so (Deletant 1989:201).

Not surprisingly, Russians in the MSSR for the past forty years overwhelmingly have considered Russian their mother tongue (98-99%). More interesting, however, are the statistics on Moldavian and Russian as "second" languages. Whereas as of 1979, 46% of all ethnic Moldavians claim Russian as their second language, only 10.6% of the Russians say the same about Moldavian (Deletant 1989:203, 205). These statistics clearly show a desire to acquire Russian on the part of the Moldavians (obviously for practical and career-furthering purposes), and, at the same time, an unwillingness to acquire Moldavian on the part of the Russian speakers. With the Gypsies as exceptions in 1970 (38%), all non-Moldavian ethnic groups who have chosen a language other than their own ethnic language as their mother tongue have held strongly to Russian. The Gypsies, however, chose Moldavian.⁷ Assimilation to Russian of certain ethnic groups—Ukrainians, Jews, Belorussians, for example—is particularly high for a number of reasons. The Ukrainians, as a case in point, find themselves socially and culturally detached from their ethnic population and thus turn to the Russians, their Slavic brethren, for a sense of culture and society. The Jews are

⁷The figures on Gypsy are so low as to be suspect, however.

concentrated in urban areas, where, as I have mentioned, Russian culture—and concomitantly, language—are considerably stronger. Russian's power and prestige in the Soviet Union are obvious everywhere. Had such a declaration of Moldavian as the republic's official language actually been implemented, it would have wreaked havoc among the Russian-speaking population—those claiming Russian as their first language, as well as those who have acquired Russian for its practical applications.

5.2. The Strength of Moldavian. Moldavian (Romanian) has remained particularly strong as the local language of this region of the Soviet Union. Moldavian occupies a prominent place in the village community, where people hold tightly to tradition *and* language; in religious life; in primary and secondary education (as of 1946, the system supports the parallel teaching of Moldavian and Russian);⁸ in the media (programs in Moldavian are broadcast four-five hours daily); and in new publications. In 1985, of 1,523 book titles published, 514 (33.7%) were published in Moldavian, 17 of 53 journals (32.1%) were published in Moldavian, and 80 of 191 newspapers (41.9%).⁹ In higher education, the situation is quite different. There, Moldavian lacks such popularity, primarily because of a perceived shortfall of practical applications for the language. Although statistics vary, many university courses are taught only in Russian, and, in general, students enter programs in Russian, not Moldavian studies. In 1982, for example, 100 students entered the Department of Philology at the University of Kishinev to study Russian language and literature, only fifty to study Moldavian language and literature (Deletant 1989:207).

There appears to be little danger of Moldavians russifying. All indications are that Moldavian remains very strong, particularly in rural areas. Russification currently is taking place among the Ukrainians, Bulgarians, Jews, and other minority ethnic groups in the republic, large numbers of whom list Russian as their second, and some even as their primary language.

6. Conclusion. In just four years (from 1987 to 1991), the Moldavian

⁸As of 1970-71, however, figures showed a decline in the number of Moldavian language teachers.

⁹Figures such as these no doubt have become more heavily weighted toward Moldavian/Romanian over the course of the last two years.

language situation has changed dramatically. An interesting dilemma—potentially a very explosive one—lies ahead for Moldavian. According to Michael Bruchis:

В отношении языка коренного населения республики Молдова ... [д]остановлением о статусе русского как официального языка на всей территории СССР, Горбачев узаконил фактически дальнейшее неравноправное положение языков нерусских народов советской империи. Если Молдова останется в подчинении Москвы, разрыв между литературной и разговорной формами языка восточнорусских (бессарабских) румын ... будет неминуемо расти и углубляться. 'Concerning the language of the core population of the republic of Moldova[,] ... by giving Russian status as the official language of the entire territory of the USSR, Gorbachev in fact has legislated a more unjust situation for the languages of the non-Russian peoples of the Soviet empire. If Moldavian remains under Moscow's control, the split between the literary and colloquial forms of the language of the east Russian (Bessarabian) Romanians ... will grow and deepen uninterruptedly.'

The concern that colloquial Moldavian will grow more distinct from the literary norm is a legitimate concern, but one quite different from concern of the early 1980s that henceforth it always would be characterized as a language distinct from Romanian. In essence, one linguistic concern has replaced another, and because of the nationalist spirit aroused by the Moldavian language itself, it seems that any "official" status for Moldavian—whether or not it is *ever* recognized as a republic language—is in jeopardy. Nowhere else in the Soviet Union are an ethnic group's national character and pride so tied to the language that they collectively speak. The battle lines have been drawn by ethnic Moldavians, and the language that they speak is being used as a weapon against the central government in Moscow. If the current situation in the MSSR is allowed to continue, Moldavian speakers will be facing not only having lost the battle to have Moldavian declared the republic's official language, but also the prospect of being forced to accept Russian as the language of use in all official regional functions. If statistics are merged, we see that in 1970, 2,438,896 (68.3%) of the population "knew" Moldavian, while 1,804,151 (50.5%) knew Russian. By 1979, 2,612,944 (66.1%) had a knowledge of Moldavian and 2,458,853 (62.2%) a knowledge of Russian (Deletant 1989:205-206).

With an overbearing—and ever-increasing—Russian-language presence in the Moldavian Soviet Socialist Republic, and a continued lack of attention to the actual form of spoken Moldavian, Bruchis' words will indeed ring true.

REFERENCES

- Argumenty i fakty*: "Moldova: v poiskah svoego puti," No. 33/514/, August 18-24, 1990, p. 4. "V pylu žarkih sporov," No. 28/509/, July 14-20, 1990, p. 4. "Peremeny dolžny sozret'," No. 16/497/, April 21-27, 1990, p. 2.
- Bruchis, Michael. 1988. *The USSR: Language and Realities (Nations, Leaders and Scholars)*. Columbia University Press, New York.
- _____. 1987. "The Language Policy of the Soviet Communist Party: Comments and Observations." *East European Quarterly*, XXI:2:231-257.
- _____. 1984a. "The Language Policy of the CPSU and the Linguistic Situation in Soviet Moldavia." *Soviet Studies*, XXXVI:1:108-126.
- _____. 1984b. *Nations, Nationalities, People: A Study of the Nationalities and Policy of the Communist Party in Soviet Moldavia*. Columbia University Press, New York.
- _____. 1983. "National History—Ancient and Modern—as Presented in the Moldavian Soviet Encyclopedia." *Crossroads*, X:165-195.
- _____. 1982. *One Step Back, Two Steps Forward: On the Language Policy of the Communist Party of the Soviet Union in the National Republics. (Moldavian: A Look Back, a Survey, and Perspectives, 1924-1980)*. Columbia University Press, New York.
- _____. 1980. "Pratiques et théories de la traduction en URSS: Causes et effets dans l'interaction des langues des peuples soviétiques dans la période d'après-guerre." *Nationalities Papers*, VIII:2:147-178.
- _____. 1979. *Rossija, Rumynija i Bessarabija (1812—1918—1924—1940)*, The Russian and East European Research Center, Tel-Aviv University.
- Chicago Tribune*: "Language Law Passes Key Test in Moldavia," August 30, 1989, p. 5(1). "250,000 Moldavians Rally for Own Language," August 28, 1989, p. 5(1).
- Crowther, William. To appear. "Moldova in the Soviet Political Context," unpublished manuscript.
- Deletant, Dennis. 1989. "Language Policy and Linguistic Trends in Soviet Moldavia." *Language Planning in the Soviet Union*, ed. by Michael Kirkwood. Macmillan Press, University of London, pp. 189-215.
- Dyer, Donald L. To appear. "Moldavian Linguistic Realities." *Papers from the*

- Fourth International Conference on the Non-Slavic Languages of the USSR*, ed. by Howard I. Aronson. Columbus, Ohio.
- _____. 1990a. "Moldavian, Part 2: Korlëtjanu's Gambit" *The Non-Slavic Languages of the USSR: Linguistic Studies*, ed. by Howard I. Aronson, Chicago, pp. 93-105.
- _____. 1990b. "Moldavia and Moldavian." Paper read at the annual meeting of the American Association for the Advancement of Slavic Studies (Washington, D.C.).
- _____. 1989. "Moldavian, Part 3: How Russian Is the Syntax?" Paper read at the Sixth International Conference on the Non-Slavic Languages of the USSR (Chicago, Illinois).
- Dyru, A.M. and A.I. Čobanu. 1970. *Limba moldovenjaskê literarê kontemporanê, v. II, Fonetika și morfologija*. Kišinev (v. I, *Leksikologija*, ed. by N.G. Korlëtjanu, 1969, Kišinev).
- Glasul Națiunii*: "Declarația de la Chișinău" July 4, 1991.
- Gramatica limbii române, I, Fonetica și morfologija*. 1963. Bucharest.
- Il'jašenko, T.P. 1983. *Formirovanie romanskix literaturnyx jazykov. Moldavskij jazyk*. Kišinev.
- Jackson Clarion-Ledger: "Thousands of Moldavians March for Independence," December 17, 1990. "Youths Riot after Funeral in Moldavia," November 17, 1990. "6 Killed, 30 Hurt as Moldavian Ethnic Unrest Mounts," November 3, 1990, p. 7(A). "Unrest Imperils Soviet Republic," October 27, 1990, p. 3(A).
- Kazazis, Kostas. 1982. "How Non-Rumanian Is 'Moldavian'?" *Papers from the Second Conference on the Non-Slavic Languages of the USSR*, ed. by Howard I. Aronson and Bill J. Darden, Columbus, Ohio, pp. 224-229.
- Korlëtjanu, N.G. 1983. *Moldavskij jazyk segodnja*. Kišinev.
- Magdan, L. 1929. *Gramatica moldovenească*. Tiraspol.
- Memphis Commercial-Appeal*: "Soviet Moldova Republic Begins Talks with Rebellious Ethnic Group," John-Thor Dahlburg, October 28, 1990.
- New York Times*: "Language Stirs Pride and Anger in Moldavia," Celestine Bohlen, October 7, 1990. "Ethnic Enmity Governs a New Soviet Republic," Celestine Bohlen, September 15, 1990. "Confession in Moldavia: They Speak Moldavian," Bill Keller, February 25, 1989, p. 4(A).
- Sergievskij, M.V. 1939. "Moldavskie ètjudy." *Trudy Moskovskogo Instituta Istorii, Filosofii i Literatury*, V:175-210.
- Tagliavani, C. 1958. "Una nuova lingua letteraria romanza? Il moldavo." *Actes du VII-ème Congrès International de Linguistique et Philologie Romanes*, pp. 445-452.
- Talmackaja, M.F. 1978. *Moldavskoe jazykoznanie v sovetskoj pečati*. Kišinev.

Udler, R.Ja. 1968. *Atlasul lingvistik moldovenesk, partja I și II*. Kișinev.

Weigand, Gustav. 1909. *Linguistischer Atlas des daco-rumänischen Sprachgebietes*. Leipzig.

GA IN LAK AND THE THREE "THERE" S: DEIXIS AND MARKEDNESS IN DAGHESTAN¹

Victor A. Friedman
University of Chicago

As Nazarov (1974:22) pointed out, Northeast Caucasian deictic systems are still in need of detailed study. The five-term series of Lak, while not the most complicated system of Daghestanian deictics (cf., e.g., Abercrombie 1889, Charachidzé 1981), is in need of more attention than it has received, particularly because the descriptions in the available grammars do not reflect the current state of the language, nor do they give a sense of the markedness relations obtaining among the five. In this paper, I shall attempt to demonstrate which of the five Lak deictics is the least marked and also determine the features for which the remaining terms are marked.

Lak is one of the five literary languages of the Daghestanian group in the Northeast Caucasian family and is spoken by close to 100,000 people in southern Daghestan (Ibragimov 1991:111). Its five-term system of demonstrative pronouns and related adverbs and adjectives is usually described as distinguishing two marked degrees of height with relation to the speaker in addition to a three-way opposition based on relation to speaker and addressee. The base forms are summarized in Table 1. The commonly cited definitions are given in boldface, additional meanings that have been given in some sources or may be relevant are given in parentheses.²

TABLE ONE

| | DEMONSTRATIVE- | LOCAL | MANNER |
|--|----------------|-------|--------|
| NEAR TO SPEAKER (new, following, focus) | va | ši- | ukun |
| NEAR TO ADDRESSEE (old, preceding) | mu | mi- | mukun |
| DISTANT FROM BOTH, NEUTRAL (level, opposite) | ta | ti- | tukun |
| BELOW SPEAKER (neutral) | ga | gi- | gukun |
| ABOVE SPEAKER | ka | ki- | kukun |

¹An earlier version of this paper appears in *A Calculus of Meaning: Markedness, Deixis and Distinctive Features*, ed. by Edna Andrews and Yishai Tobin. Amsterdam: Benjamins. The present version has been expanded and the conclusions revised.

²Lak examples are all given in Latin transcription of Cyrillic Literary Lak orthography.

Demonstratives decline after the pattern of *ga* in Table Two. (But *mu* keeps its vowel in the oblique and its plural is *mij*.) The other oblique cases are formed by replacing the *-l* of the genitive, e.g., dative *gana-n*. Lak also has absolutive affixes and clitics that attach to the nominative.

TABLE TWO

| PRONOUNS | | singular | | plural | |
|---------------|---------|----------|--------|--------|-----------|
| | | nom. | gen. | nom. | gen. |
| first person | | na | ttu-1 | žu | žu-1 |
| second person | | ina | vi-1 | zu | zu-1 |
| demonstrative | class 1 | ga | gana-1 | gaj | gajinna-1 |
| | 2 | | gani-1 | | gajnnu-1 |
| | 3 | | | | |
| | 4 | | | | |

The personal pronouns, which are limited to the first and second persons, are opposed to demonstratives, any of which can fill the function of a third person, for which no special form exists.³ In keeping with the markedness hierarchy identified by Silverstein (1976), the personal pronouns remain in the nominative (absolutive) case regardless of whether they are the subject or direct object of a normal transitive or intransitive verb (examples 1a, 2a, 3a), whereas the demonstrative pronouns, like all other nominals, will be in the absolutive as the subjects of normal intransitives and as the objects of normal transitives, but will be in the genitive (which serves ergative and other functions) when the subject of a normal transitive, as in examples (1b, 2b, 3b):⁴

³It is important to note that any of the five deictics can occur alone in the function of a third person pronoun; none of them is limited to a purely demonstrative function.

⁴By *normal* I am excluding verbs and constructions that take dative or ablative subjects, which neutralize the distinction personal/nonpersonal (cf. Kibrik 1978).

2c. ttu-n va čvu b-axxan ččaj b-ur

I want to sell this horse

2d. gana-n va čvu b-axxan ččaj b-ur

He wants to sell this horse

2e. ttu-šša va čvu b-axxan bjuqlaj b-ur

I can sell this horse

2f. gana-šša va čvu b-axxan bjuqlaj b-ur

He can sell this horse

| | |
|------------------------------|--------------------|
| 1a. na naj \emptyset -ur-a | I am coming |
| 1b. ga naj \emptyset -ur | He is coming |
| 2a. na b-axxara čvu | I sell the horse |
| 2b. gana-l b-axxajčvu | He sells the horse |
| 3a. na ga ucarā | I am bringing him |
| 3b. ganal na ucarā | He is bringing me |

Morphologically, the demonstratives agree with their referent according to class as indicated in Table Two.⁵ Local deictics can either take the local case affixes directly, e.g., *ši-vu* 'in here', *ši-v-un* 'into here', *mi-vu-n-n-aj* 'towards in there by you (class 2)' or three special local affixes:

| | |
|-------|-----------|
| -kku | location |
| -kkun | toward |
| -šša | away from |

⁵In general, class 1 is used for mature male humans, class 2 for mature female humans, class 3 for other animates and some inanimates, and class 4 for inanimates and some animates. The demonstrative agreement type, whose patterns of syncope are illustrated in Table (i) as Type A, is peculiar to those pronouns and the morphologically related definite adjectives. It differs significantly from the syncope patterns in most other agreeing parts of speech (verbs, numerals, adverbials, emphatics, directive cases, durative adjectives), illustrated here as Type B. A third pattern, Type C, is found only in the oblique reflexive pronoun, while Type D occurs in possessive pronominal adjectives in *-ssa*.

Table (i)
Patterns of Syncope in Lak Class Markers

| sg pl | sg pl | sg pl | sg pl |
|------------------|------------------|------------------|------------------|
| 1 2 3 4 | 1 3 2 4 | 1 2 3 4 | 1 2 3 4 |
| TYPE A | TYPE B | TYPE C | TYPE D |

e.g., *šikku* 'here,' *kikkun* 'up thither,' etc.

Given that all of Lak's five demonstratives can and do function in roles played by third person pronouns in languages that make such a distinction, our first question can be which, if any, of the five deictics occurs most freely in this use. In the first full description of Lak, published in 1890, but based on fieldwork done in the early 1860's, Uslar (1890:63) states unequivocally that *ta* serves this function. He adds that *ta* is used when the object is undifferentiated with respect to its position in relation to the speaker and addressee, especially when it is to one side of both. He goes on to state that *va* is used for objects closer to the speaker than to the addressee, *mu* is used for those closer to the addressee than to the speaker, *ka* is used for objects higher than the speaker, and *ga* for those lower than the speaker. The next description of Lak phonology and morphology was published by Žirkov (1955). In his discussion of demonstratives (Žirkov 1955:71), he repeats Uslar's formulations, but with some modifications. Although Žirkov states that *ta* is the most neutral with respect to level and distance, he also notes that for *ga* the meaning 'lower' has been significantly weakened, and that it is often used as a neutral demonstrative or a simple third person pronoun. His formulations for *va* and *mu* are also slightly more general: *va* is defined simply as denoting the 'nearest,' while *mu* is defined as being in the 'sphere' of the addressee. For *ka* he states that the indication of the object as higher is usually quite clear. In the next Lak grammar, Murkelinskij (1971:162) repeated Uslar's formulations without any modification, and did so again in his Lak school grammar (Murkelinskij 1980:85).

For my own study, I examined deictic usage in texts in a variety of genres: literary prose, expository prose, and epic poetry. The literary prose turned out to have the greatest variety of deictic usage, although the expository prose also provided useful examples. Due to its dependence on formulas, rhythmic structure, and repetition, Lak epic poetry provides a relative paucity of demonstratives. The author or narrator is much more likely to repeat the name of the referent. The small expository prose sample of about 2,600 words was striking in that *va* and *mu* accounted for over 90% of the 62 deictics, with *ga* accounting for the remainder. This will prove useful in our discussion of the distinction between *va* and *mu* later. My main source of material, however, was a collection of folk narratives (Xalilov 1976), which

contained 540 deictics in a corpus of over 7,000 words. With one exception, the 33 narratives were collected between 1950 and 1966 from more than 15 speakers from the Kumux dialect area, which is the largest and also provides the basis for literary Lak.

From what has just been said, it can be seen that *ta* and *ga* are the two most likely candidates for unmarked demonstrative/third person pronoun. In her chapter entitled "Myths About Markedness," Andrews (1990) rightly points out that statistical rarity is not a defining characteristic of markedness. It can and does happen in language that the meaning for which a word is marked is one which speakers have frequent cause to specify, e.g., personal confirmation in Macedonian (cf. Friedman 1977). In looking at a complex deictic system in a broad range of contexts, however, if we see one set accounting for almost 52% of the examples, as is the case with Lak *ga* in our corpus of folk tales, and another totalling less than 9%, as occurs with *ka*, we can at least look first at the less frequent form for a defining characteristic. And when we discover that the deictic traditionally described as the most neutral accounts for under 3% of the examples in the same corpus, we have reason to investigate carefully the basis of the traditional description.⁶

In the folktale corpus, the 278 occurrences of forms based on *ga* do not display any invariant meaning that can be connected with the concept "lower." Examples (4) is typical:⁷

- (4) Ivkkun ur, qqaivkun ur ca šjaravu ca čjavuḱul malla ivkun ur.
Cal, njužmar qini *ga* mizitravu quran kkalaj ivkun ur. (Xalilov 1976:204)

Once upon a time in a certain aul there was a clever mullah. One Friday **he** was reading aloud the Koran in the mosque.

On the other hand, the occurrences of *ta* all display a sense of contrast or opposition, as in (5) and (6):

- (5) Žuva urča insan uru, **taj** ḱija. (Xalilov 1976:217)
We are nine people, **they** [are only] two.

⁶In the folktales, *mu* was slightly more frequent than *va*, accounting for almost 24% of the deictics, while *va* accounted for almost 19%.

⁷Cf. also example (14) below.

- (6) Ca "hura" buvkuna qquv - a's kussakssa,
 Šadšivu dullajssa lakral čuluxa,
 Ca "uh" qqak buvkuna **tamur** čuluxa,
 Paṭima paxrulij x'x'ičux bavčukun. (Xalilov 1969:33)
 A "hurrah!" like thunder and lightning was heard
 from the rejoicing Lak side
 An "ugh" was heard grumbling from **their** [the enemy's] side
 As Patima proudly rode in front of them.

Similarly, Lak fixed expressions contrasting proximal and distal local expressions use the forms based on *ti-* as in example (7)

- | | |
|--------------------------|--|
| (7) <i>tij-šij</i> | on both sides (lit. 'thereon hereon') |
| <i>tin-šin baqqa</i> | without hesitation (lit. 'thither-hither not-being') |
| <i>tixa-šixa</i> | thence and hence |
| <i>tixunmaj-šixunmaj</i> | thither and hither (class 3) |

Example (8) occurs in the context of two rooms. In one, the faithful wife entertains importunate suitors one at a time, in the other stands a kettle of boiling water into which the suitors fall when they flee her supposedly returning husband. The use of the locational deictic *ti-* not only stresses the "oppositeness" of the second room, but also conveys the sense of the victim walking in expecting a level floor before falling to his death:

- (8) **Gagu, tivun** naj una, *ḵunḵurduvun* ahnu ur. (Xalilov 1976:216)
He, too, going into there, fell into the kettle.

In view of this evidence, we may well wonder why the grammars all state or give the impression that *ta* is the least marked demonstrative and the closest equivalent to a third person pronoun. I believe the answer is to be found in a combination of time and space. Although Uslar writes that he worked in Kumux with his chief informant, Abdulla Omar-ogly, Žirkov states that the dialect described by Uslar was that of Vicxi, which is a dialect area to the north of the Kumux region. In our corpus, more than half the examples of *ta* occur in a single tale, the only one recorded before 1950 from a speaker from outside the

Kumux area, viz., in 1934 in the aul of Kurkli in the northern Vicxi region. I would suggest that while *ta* may well have been the unmarked demonstrative for Vicxi speakers in the 1860's—and this apparently reflects the original situation (cf. Burčuladze 1985)—Kumux speakers of today have shifted the marking. Markedness for depth in *ga* has been replaced by markedness for *contrast* or *opposition* in *ta*. The persistence of Uslar's description is probably due to a combination of respect for his work and lag time between normative and descriptive grammar.⁸

Unlike *ga*, *ka* has clearly retained its height orientation. All of our examples clearly refer to some sense of physical height, even when they are temporal. Thus, for example in a tale about Nadir Shah of Persia planning a campaign against Lakkia, which is physically higher than Persia, forms of *ka* are used repeatedly to refer to things connected with Lakkia, as in example (9):

- (9) Lakral a'rallunnavu turča bikanvagu, *ka* čumal šanazara insan ivkun ur. (Xalilov 1976:207)

As for the Lak army, at **that** time it had three thousand people.

Similarly, example (10) shows that *ga* can occur as a referent to a preceding item in the discourse which can also be explicitly referred to with *ka*, which here functions much as English 'the above.' This is additional evidence that *ga* is unmarked with respect to *ka*:

- (10) Podležaššee va skazuemoe calčimur daražalul člentrur. **Gaj** qanaqissar predloženijalul hanu. ...

Xxal dannu ukunssa kira predloženie: *Int durkunni*. **Va** predloženie sakin x'unu dur tak kiva calčimur daražalul člentrāja—podležaššilija va skazuemilija. **Ka** predloženie lači dannu. ... *Zuntavun ġilissa int ččjani durkunni*.

Ka predloženijaluvu calčimur daražalul člentrū—*int durkunni*—qanaqissar predloženijalul hanunu. (Buržunov 1975:27)

⁸It is also worth noting that this part of the system is unstable. In the Arakul dialect, spoken in two auls in southernmost Lakkia, *va* can substitute for any of the demonstrative forms (Xajdakov 1966:28-30). It may be that Azeri influence has contributed to this breakdown.

The subject and predicate are the primary constituents. They constitute the basis of the sentence. ...

Look at such a two-member sentence: *Spring arrived*. This sentence is made up of only two primary constituents: a subject and a predicate. Lengthen this sentence. ... *A warm spring came early to the mountains*.

In **this** sentence the primary constituents—*spring arrived*—constitute the basis of the sentence.

Thus, if we take the subset of Lak demonstratives that have traditionally been described as referring to height with relation to speaker (above-level-below), and which we can call the dimensional set, we find the following situation. In the recent past and in the North, unmarked *ta* was opposed to *ka* marked for 'highness' and *ga* marked for 'lowness,' and *ta* had a chief contextual variant meaning of 'opposite.' Over time, however, 'opposite' has developed into the marked meaning of *ta* while *ga* has lost its marking for 'lowness' and become the unmarked member of the opposition.

The next question is how *ga* relates to *va* and *mu* and how these two relate to each other. In view of traditional descriptions, we can refer to this group for the time being as the personal set. Here again Žirkov (1955:71) seems closer to a broader generalization than Uslar. It is clear that "closer to the addressee" will not account for the use of *mu* in example (11) but "in the sphere of the addressee" will, if "sphere" is understood in a sense of "sphere of interest":

(11) Calčinmanal kunu bur: - qqarššunni cukunčav ššinavun. Ttun
mu ttula jarunnin kkavkkunni, - kunu. (Xalilov 1976:214)

The first one said: "No way did it touch the water. I saw it with these my own eyes."

Thus, while the speaker's eyes are obviously closer to him than to his interlocutor, he uses *mu* because he wishes to draw the addressee's attention to his eyes.⁹ Consider also the shift in deictic reference in

⁹The Lak deictic *mu* is reminiscent of Turkish *şu*, which like Lak *mu* is described as marked for closeness to the addressee, but which appears to be marked for deictic salience (Ščerbak 1977:129, cf. also Anderson and Keenan 1985:285),

example (12) in which a rich merchant is giving deathbed instructions to his spendthrift son:

- (12) Ina za jadajssa aqqara, qinu xarž dullalissa, jaluv qqaavçussa ura. Ttul qus vin **mukunssa** o'rmulun dijal qqax'unnissar. **Vaj** vila dustural inava lijan uvnu, ciçar daqqa livčukun, **gaj** vija kura bajannissar. **Gajnnal** inava qaivtukun, vin o'rmu qqaččan x'unnissar. **Mu** čumal ina ca x'x'agu lavsun žula qqatluł kılčınmur uttussanuj a's x'u. (Xalilov 1976:208)

You do not save anything, you are a spendthrift and a wastrel. My wealth is not enough for **this kind** (mu) of life you lead. **These** (va) friends of yours having ruined you and left you with nothing, **they** (ga) will turn their backs on you. When **they** (ga) leave you, you will not want to live. At **that** (mu) time, take a rope and hang yourself from the second beam of our house.

The use of *mukunssa* by the father to refer to the son's life style is in keeping with second person reference, but it also refers to the topic introduced in the sentence which preceded it. The use of *vaj* in the following sentence, however, is not due to the son's friends being closer to the father but rather to their being introduced as a new topic related to the theme of the discourse; they had not been referred to at any previous point in the narrative. This done, they are indicated by the neutral *gaj* in subsequent clauses. In the final sentence, the father is again referring anaphorically to what he has just said; hence, his choice of *mu*. In fact, the expression "at that time" usually refers to the time of an event just referred to, and hence its standard expression in Lak is *mu čumal*. However, if the narrator wishes to emphasize the gap in time between two events, then one of the other deictics will be used, as in example (9) cited above or example (13), in which the speaker, who has not seen his interlocutor for 15 years, is referring to the fight they got into the last time they met:

- (13) Daķnijriv žula aķmaqssa bijavu. Amma ci turčagu, ga čumal culķlul maķ nexxajn ššuna. (Xalilov 1976:214)

rather than any sort of proximity. Unlike *šu*, however, *mu* is used anaphorically, whereas *šu* is used to introduce following material, the less emphatic but more proximate *bu* being the deictic of anaphora.

Truly we were fools. But anyway, **that** time the fox's tail did touch the river.

Example (14) provides a compact contrast of the three deictics that are not marked for dimension:

- (14) **Mugu** maḡattal x'unu dur **va** iširaj **va** **ganin** qast x'unu dur insantural ṭimur mjaḡžanssa buriv xxal ban. (Xalilov 1976:204)
She, too, was amazed at **this** thing, and and **she** decided to see if what people were saying was true.

The woman was introduced into the narrative in the preceding sentence as the wife of a *mutalim* 'seminarian,' and *mu* is the first anaphoric reference to her. The *thing* she was amazed at was the focus of the first main incident of the story. Thus *va* is something new to her but known to the reader/listener. It is also further back in the narrative. The third demonstrative, the class 3-4 dative singular of *ga*, refers again to the woman, whose pronominal identity was established by *mu* and which fact therefore now shifts to the background.

Consider in this context also the use of *va* in example (10) above, repeated here as example (15):

- (15) Podležaššee va skazuemoe calčimur daražalul člentrur. **Gaj** qanaqissar predloženíjalul hanu. ...
 Xxal dannu ukunssa kira predloženie: *Int durkunni*. **Va** predloženie sakin x'unu dur tak kiva calčimur daražalul člentraja —podležaššilija va skazuemilija. **Ḳa** predloženie lači dannu. ...
Zunttavun ġilissa int ččjani durkunni.
Ḳa predloženíjaluvu calčimur daražalul člentru—*int durkunni* —qanaqissar predloženíjalul hanunu. (Buržunov 1975:27)

The subject and predicate are the primary constituents. **They** constitute the basis of the sentence. ...

Look at such a two-member sentence: *Spring arrived*. **This** sentence is made up of only two primary constituents: a subject and a predicate. Lengthen **this** sentence. ... *A warm spring came early to the mountains*.

In **this** sentence the primary constituents—*spring arrived*—

constitute the basis of the sentence.

The author began the paragraph with an exposition on the primary and secondary constituents of a sentence, with all demonstratives being of the *ga* type. He then begins a new paragraph telling the reader to consider a concrete example. In referring to this newly introduced material, he uses *va* on the first reference and subsequently switches to the marked positional deictic *ka* in further references.

In an expository text such as the introduction to Xajdakov and Žirkov (1962:8-10, 16-20) the reader is led from point to point with no sense of digression. In this text of some 2,600 words there are 45 occurrences of deictics based on *mu*, 12 from the *va* group, 5 based on *ga*, and none based on *ta* or *ka*. The explanation for this is the fact that the majority of deictics in this text refer to preceding material that is at the same time not newly introduced. The examples with *va* refer to new or following material, while the few with *ga* are neutral with respect to these values. There are no binary contrastive contexts, hence the lack of *ta*, and likewise no references to a specific foregoing item (as in example 10/15), hence the lack of *ka*. The following examples are illustrative of these contrasts:

- (16) Amma ɣaɣinussa ɣininin cavagu, ɕanssa bux'urɕagu ššallussa — ja ɣivissa, ja qunmassa lakku mazral va o'rus mazral slovar' bukkann qqabuvssija. **Va** ɕanassa slovar' qanaqissar calɕinssa **mukunssa** slovar'nu. (Xajdakov and Žirkov 1962:9)

But until today not one more or less satisfactory Lak-Russian dictionary has been published. This (*va*) present dictionary constitutes the first such (*mu*) dictionary.

- (17) ɕardaššalssa soɕetaniarttu **ukun** kkakkan buvnu bussar. (Xajdakov and Žirkov 1962:17)

Nominal collocations are shown thusly (*va*):

- (18) Lakku mazral va o'rus mazral slovar' qanaqissar ɕjavuɕil mjunpat lasunssa lunu. **Munil** qunmassa kumag bantissar lakral školardal uɕitel'turan, ... (Xajdakov and Žirkov 1962:10)

The Lak-Russian dictionary is a book useful for everyone. It (*mu*) will be of great help to teachers in Lak schools, ...

- (19) Muqul ca mjaʔna **gamur** maʔnalijatu ʔunt x'iriv bivx'ussa baši ʔanssa a'rab tarixraj ličij durnu dussar; ... (Xajdakov and Žirkov 1962:17)

A bold-faced Arabic numeral followed by a period differentiates one meaning of a word from another (*ga*) meaning.

Examples (16), (17), and (18) show contrasts of various forms based on *va* and *mu*: the *va* forms refer either to a new topic or to following material, whereas the *mu* forms refer to the immediately preceding clause. Additional evidence for the meanings of *mu* and *va* can be found in the deictics most frequently used in certain expressions such as *munixlunu* 'therefore,' *mukunma* 'likewise,' which typically refer to some immediately preceding statement or item (see also the discussion of *mu ʕumal* 'at that time' above), whereas expressions such as *vaj maqssa šinardij* 'in recent years,' *va maqssa ʕumuvu* 'recently' are used to introduce a new topic in the discourse. Another common usage is an adjectival form of *ga* to mean 'another,' as in example (19). This usage is conditioned by the fact that as the least marked deictic, *ga* is used to qualify an item that is being separated from but not contrasted with some other.

In terms of markedness relationships, *va* and *mu* form a pair opposed to *ta* and *ga*, which are also paired. The four together are all opposed to *ka*, which is clearly marked for the feature 'height.' I would argue that the opposition between *va/mu* on the one hand and *ta/ga* on the other is that the former are marked for a higher degree of discourse salience (cf. Hanks 1989, Kirsner 1979). Both *va* and *mu* are clearly associated with deictic and discourse meanings such as 'speaker/addressee,' 'anticipation/anaphora,' while both *ta* and *ga* do not mark such connections in a narrative. In a sense, *va/mu* are more 'demonstrative' than *ta/ga*, but it would be a mistake to use 'demonstrative' as a distinctive feature in view of the fact that all five deictics can function as both demonstratives and as personal pronouns; they can all modify another noun phrase or stand alone. But it is clear from the foregoing examples that *va/mu* direct the addressee's or reader's attention to specific places in the text/discourse such as 'sphere of speaker, following information, new information' on the one hand and 'sphere of addressee, preceding information, old information,' on the other. The deictics based on *ta* and *ga* do not have this type of discourse

function.

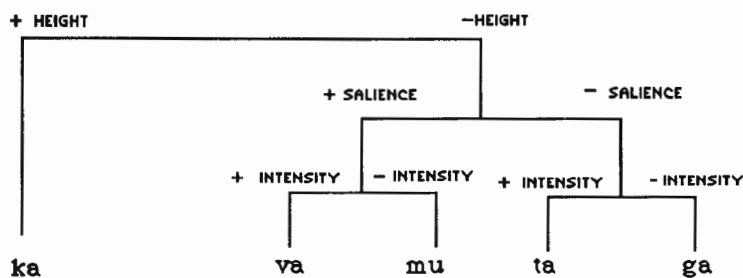
Within each pair, it is clear that *ta* is marked with respect to *ga*, and I would argue that *va* is marked with respect to *mu*. As noted above, frequency of occurrence does not constitute proof of markedness. Nonetheless, the fact that *ga* in narrative prose and *mu* in expository prose are the overwhelmingly predominant pronouns suggests that it is the other pronouns that can be defined in terms of an invariant meaning. In the case of *ta*, a meaning such as 'contrast' or 'opposition' was suggested above, while for *va* the meaning is associated with 'anticipation, speaker, new topic.' Kirsner (1979) has identified a scale of deictic emphasis in Dutch that is relevant for Lak. According to such a scale, deictics are marked for the emphasis they place on the act of deixis itself (high vs low intensity) rather than on a feature such as relative proximity. This seems to be particularly applicable to *va* and *ta*. In the case of *ta* marking for intensity without marking for salience gives the meaning 'opposite, contrasting,' while in the case of *va* marking for both salience and intensity gives 'speaker, anticipation, new topic.' For *mu*, marking for salience without marking for intensity yields 'addressee, preceding, anaphora,' while, *ka* marks maximal distance and *ga* is the most neutral deictic.

These relations are given as a feature matrix in Table Three and as a hierarchical diagram in Figure One.

TABLE THREE

| | height | salience | intensity |
|-----------|--------|----------|-----------|
| <i>ka</i> | + | 0 | 0 |
| <i>ta</i> | - | - | + |
| <i>ga</i> | - | - | - |
| <i>va</i> | - | + | + |
| <i>mu</i> | - | + | - |

FIGURE ONE



The foregoing analysis differs significantly from previous analyses in two important points. First, it has shown that *ga* and not *ta* is the unmarked deictic in modern literary Lak. This appears to be the result of a relatively recent shift through which *ta* has become restricted to its chief contextual variant meaning of 'oppositeness' while *ga* has lost its specificity for 'lowness.' Second, in differentiating *va*, *mu*, and *ta* the analysis has shown that relation to speaker and addressee is not the defining factor. With regard to what do in fact constitute the semantic features, I have suggested a combination of markings for 'salience' and 'intensity' to account for the differences in meaning.

REFERENCES

- Abercrombie, John. 1889. *A Trip through the Eastern Caucasus: With a Chapter on the Languages of the Country*. London: E. Stanford.
- Anderson, S. and E. Keenan. 1985. "Deixis." *Language Typology and Syntactic Description*, vol. III, ed. by T. Shopen, 259-308. Cambridge: Cambridge.
- Andrews, Edna. 1990. *Markedness Theory: The Union of Asymmetry and Semiosis in Language*. Durham: Duke University.
- Burčuladze, G. T. 1985. Čvenebit nacvalsaxelta pujeebi lakurši. *Iberiul kavkasiuri enatmecniereba* 23.301-07.
- Buržunov, H. H. 1975. *Lakku maz: 4 klassranssa lu* (3rd edition). Maxačkala: Dagučpedgiz.
- Charachidzé, Georges. 1981. *Grammaire de la langue avar*. Paris: Jean-Favard.
- Friedman, Victor A. 1977. *The Grammatical Categories of the Macedonian Indicative*. Columbus: Slavica.
- Hanks, William F. 1989. "The Indexical Ground of Deictic Reference." *CLS 25, Part 2: Parassession on Language and Context*, ed. by B. Music et al., 104-22 Chicago: Chicago Linguistic Society.

- Ibragimov, M.-R. A. 1991. "Narody Dagestana v XX v. (ėtnografičeskie problemy)." *Rasy i narody* 21:100-127.
- Kibrik, A. E. 1978. *Materialy k tipologii ěrgativnosti. 2. Laksij jazyk. (Problemnaja grupa po ěksperimental'noj i prikladnoj lingvistike, Predvaritel'nye publikacii 12.7)*. Moscow: AN SSSR, Institut russkogo jazyka.
- Kirsner, Robert S. 1979. "Deixis in Discourse: An Exploratory quantitative Study of the Modern Dutch Demonstrative Adjectives." *Discourse and Syntax*, ed. by Talmy Givon, 355-75. New York: Academic Press.
- Murkelinskij, G. B. 1971. *Grammatika laskogo jazyka*. Maxačkala: Dagučpedgiz.
- _____. 1980. *Lakku maz: 5-6 kalssirtanssa učebnik*. Maxačkala: Dagučpedgiz.
- Nazarov, V. N. 1974. *Razyskanija v oblasti istoričeskoj morfologii vostočnokavkazskix jazykov*. Maxačkala: Dagučpedgiz.
- Ščerbak, A. M. 1977. *Očerki po sravnitel'noj morfologii tjurkskix jazykov: Imja*. Moscow: Nauka.
- Silverstein, Michael. 1976. "Hierarchy of Features and Ergativity." *Grammatical Categories in Australian Languages*, ed. by R. M. Dixon, pp. 112-71. New Jersey: Humanities Press.
- Uslar, P. K. 1890. *Laksij jazyk. (Ėtnografija Kavkaza. Jazykoznanie. 4)*. Tbilisi: Upravlenie Kavkazskogo Učebnogo Okrug.
- Xajdakov, S. M. 1966. *Očerki po laskoj dialektologii*. Moscow: Nauka.
- _____. and L. I. Žirkov. 1962. *Lakku mazral va o'rus mazral slovar'*. Moscow: AN SSSR.
- Xalilov, X. M., ed. 1969. *Lakral laqi balajrdu*. Maxačkala: AN SSSR, Dagestanskij filial.
- _____, ed. 1976. *Satira i humor narodov Dagestana*. Maxačkala: AN SSSR, Dagestanskij filial.
- Žirkov, L. I. 1955. *Laksij jazyk*. Moscow: AN SSSR.

THE MORDVIN MASTORAVA ('MOTHERLAND') MOVEMENT IN SOCIOHISTORICAL PERSPECTIVE*

Jeffrey Harlig
Indiana University

1. Introduction. Studying the sociolinguistics of the Uralic languages is nothing if not an exciting experience: While the languages go on largely unperturbed, the societies in which they are spoken insist on changing dramatically at a moment's notice. Hungary became democratic almost overnight in 1989; Estonia declared independence in 1990, and in September 1991 became a full member of the United Nations. In the summer of 1991 I had the unique opportunity to visit the Mari Republic, titular republic of a Volga-Finnic ethnic group, and get an idea of language maintenance problems there. Within ten days of my return home, the August coup had taken place and failed in the Soviet Union, leading quickly to the end of Communist Party power and the dissolution of the Soviet Union. This startling change in the once rock-solid backdrop of all sociological, economic, and historical studies of the Soviet Union—i.e., the disappearance of the monolithic communist state—has sent social scientists scrambling. For the Uralic speakers of the former Soviet Union, and all other minority-language speakers, the end of centralized social planning and a change in status from a "Soviet nationality" should mean renewed hope and possibilities for ethnic revival and language maintenance. But what does it mean, in fact? With the exception of Estonia, every eponymous Uralic administrative unit of the former Soviet Union lies within the Russian Federation, and the vast majority of the Uralic peoples for whom those districts are named live in Russia as well. For these peoples, as for the speakers of various Turkic, Paleo-Siberian, Indo-European, and other languages, their place within a huge nation, in the company of a large majority of Russian speakers, has not changed at all. In the current paper I want to present the situation of one group of speakers—the speakers of the Mordvin languages—including not only the material

I would like to thank Allen Frank and Gregory Rogers for many helpful discussions, and Toivo Raun, Bill Fierman, and Isabelle Kreindler for valuable comments on an earlier draft of this paper. For the most part I have retained the temporal reference point of when this paper was originally submitted, that is, about mid-1992. Needless to say, many situations have changed, while some of my predictions have been borne out and others not.

presented at the 1991 Non-Slavic Languages Conference, but also a historical and social perspective in which to view the present, and future, situation of this group. The additional perspective is required by two facts. First, the *Mastorava* movement—a political, ethnic, and linguistic revival movement among the Mordvins which I reported on at the conference—appears to have stalled, and consequently, the “progress report” perspective of the original paper is no longer valid. Second, many of the proposals and declarations of the movement were meant to explicitly oppose the course of Soviet history and the policies of the then-current Soviet regime. The complete change in the internal status and external relationships of the *Mastorava* movement requires a substantial change in presentation, one that goes beyond *Mastorava* and looks at the possible future of Mordvin and other languages in post-Soviet Russia.

2.1. Ethnic divisions among the Mordvins. The name “Mordvin” is an externally imposed ethnonym referring to the two descendant parts of a once unitary ethnic group belonging to the Volga-Finnic branch of the Finno-Ugric peoples. The Mordvin peoples have been distinct since the sixth century A.D. (Kreindler 1985b): the Moksha group lives primarily in the southwest and the Erzia group primarily in the northeast of the range of Mordvin habitation. The two groups have been influenced by different outside contacts, resulting in substantial differences in customs, language, and even physical appearance. A proverb among the Mordvins refers to the two groups as “a people with one head and two mouths” (Kreindler 1985b:238). For convenience, and by tradition, the Mokshas and Erzias will be referred to as Mordvins in this paper, except where differences between the two language groups are at issue.

2.2. Location and demographics of the Mordvins. The Mordvin (formerly “Autonomous”) Republic is located between the Oka and Sura rivers, southwest of the bend of the Volga River. According to Hajdú (1980), the Mordvins have the most highly developed agricultural and animal husbandry techniques of any Uralic group in the former USSR. There were approximately 1,154,000 Mordvins reported in the 1989 census, making them the largest Uralic group after the Hungarians and Finns. However, they are one of the most greatly dispersed of the Finno-Ugric peoples. More than 70% of the persons declaring themselves to be Mordvin live outside of the Mordvin Republic. The

areas of greatest habitation outside of the Mordvin Republic (in descending order) are the Kuibyshev, Penza, Orenburg, Ulyanovsk, and Gorky oblasts of the Russian Federation, and the Tatar and Bashkir Republics (Lallukka 1990). In addition, the Mordvins are in the minority in their own republic. Although the number of residents in the Mordvin Republic in 1989—964,000—was close to the total population of Mordvins in the USSR, in fact only 32.5% (about 313,300) of the republic's population were ethnic Mordvins. Sixty-one per cent were Russian and the rest (~7%) were Tatars and other nationalities. There has been a steady decrease in the number of people declaring themselves as Mordvin in census counts since 1939. This can be seen in table 1, which demonstrates that the Mordvins are doing much more poorly in maintaining their ethnic identity than the other large Finnic groups in the former USSR.

TABLE 1
Major Finno-Ugric Nationalities in the USSR (in thousands)

| | 1926 | 1939 | 1959 | 1970 | 1979 | 1989 | change 1926-89 |
|-------------|---------|---------|---------|---------|---------|---------|-------------------|
| Mordvins | 1,340.4 | 1,456.3 | 1,285.1 | 1,262.7 | 1,191.8 | 1,154.0 | -14% |
| Estonians | — | — | 988.6 | 1,007.4 | 1,019.0 | 1,026.6 | +3.8%* |
| Udmurts | 504.2 | 606.3 | 624.8 | 704.3 | 713.7 | 746.8 | +48% |
| Maris | 428.2 | 481.6 | 504.2 | 598.6 | 622.0 | 670.9 | +57% |
| Komi | | | | | | | |
| (-Permyaks) | 375.9 | 422.3 | 430.9 | 475.3 | 477.5 | 496.6 | +32% |

*Change since 1959, the first census reflecting current Estonian borders

Sources: Kreindler (1985a:43) (1926-79); 1989 figures from census data

Kreindler sees the steady decrease in the Mordvin population as "clearly a case of assimilation" since the Mordvins "have one of the lowest urbanization rates and, predictably, enjoy a fairly high fertility rate" (1985a:44). The fact that the majority of Mordvins live in areas of low Mordvin population density outside the Mordvin Republic makes them particularly vulnerable to assimilation. It must be acknowledged, however, that assimilation is also taking place within the Mordvin Republic.

3.1. Past history of Slavic contacts. It is not necessary to go into detail about the history of the Mordvin contact with Russia, as this is treated admirably by Kreindler (1985a, 1985b) and Lallukka (1990). It is sufficient to quickly summarize their discussion of this contact, to establish the background for the current condition of the Mordvins and their languages.

The Mordvins are the earliest of the surviving Volga nationalities to have made contact with Slavic peoples. Large-scale contact with Russians goes back to at least the 12th century, but archeological evidence suggests that Slavic contacts go back to the 6th century. From the 17th century on, the Mordvins struggled constantly against Russian expropriation of lands and forced conversion to Russian Orthodoxy. Under continuing pressure of Russian advances into their lands, the Mordvins were forced to flee eastward, resulting in the wide dispersal pattern evident today. By the turn of the 17th and 18th centuries the Mordvins were a minority in their own homeland. 17th- and 18th-century sources already speak of the Mordvins as heavily Russified in religion, customs, and language. Even late 19th-century observers predicted the imminent assimilation of the Mordvins and the disappearance of their language and culture.

The process of outmigration continues today. Between 1970 and 1989, the population of the republic decreased by more than 65,000—over 6% of the 1970 population of approximately 1,029,600 people. The Resolution of the First *Mastorava* Congress, which will be discussed in detail below, focuses explicit attention on emigration: “Planned ‘mistakes’ in social and economic policy have elicited an unparalleled wave of migration of the indigenous population. In only the last five years twenty-five thousand people have left the republic; of these twenty thousand have been Erzias and Mokshas.... [Local authorities] are afraid of responsibility for the evil which has been committed...” (Rezoljucija 1990:9¹).

3.2. Fate of the Finnic intelligentsia under Stalinism. While the Finno-Ugric nationalities probably did not suffer greater repression overall in the 1930s than other small peoples within Russia, a combination of historical factors resulted in great vulnerability for the

¹Page references for the Resolution refer to the materials in Moksha; the English-language renderings are from a translation by Allen Frank of Indiana University (Frank 1992).

Finnic peoples in particular. As a result of exhortations by Lenin, intellectual leaders of the Finnic nationalities sought to improve the cultural stature of their peoples by many means, but one of the main strategies was identification with the major political and cultural national entity of the Finnic group, Finland.

However, in the 1920s and 1930s, the Soviet Union came to view as a serious threat certain Finnish movements coveting parts of the Soviet Union where Finnish/Finnic speakers resided, particularly East Karelia. Stalin seized upon the existence of various nationalist, right-wing organizations in Finland to launch a propaganda campaign against that country.² At the same time, the ascendance of Marr's alternative to generally held theories of genetic relationship among languages called the Mordvin-Finnic relationship into question (Kreindler 1985b). Consequently, just as the "pan-Finnic" movement was reaching its height in the Soviet Union, association with Finland became dangerous and politically debilitating. According to Kreindler, "pan-Finnism was a very dangerous accusation" in the 1930s (Kreindler 1985b:249). Many Mordvin scholars and intellectuals, as well as the intelligentsia of the Veps (Strogaľnikova 1990), Komis (Popov 1990), Maris (Sanukov 1991), and numerous other Finnic groups, were purged during the mid-1930s for "bourgeois nationalism," and for making use of the work of "fascist" Finnish scholars (Kreindler 1985b).³ A similar fate befell Mordvin scholars at the same period of time. These activities, of course, did not prevent people from speaking their native language, but served once again to limit the public functions of language and cut off the further development of literary languages and literary traditions.

4. Language education and language loyalty among the Mordvins. Standardization and literacy campaigns for both Moksha and Erzia

²My thanks to Toivo Raun for help with these historical facts.

³For example, organizations were formed in the 1920s and 1930s to study Komi culture and history on the basis of Lenin's policies for cultural renaissance. The authors who had responded to Lenin's call were purged as nationalists in the Stalinist era. In June and July 1935 a campaign was begun to remove what they had created—now labelled "nationalist and counterrevolutionary literature"—from libraries, schools, and archives (Popov 1990). Similarly, a Veps alphabet was devised in 1931, instruction in native language began in the Leningrad district in 1932, and Veps textbooks were written. All activity halted abruptly in 1937 because of Stalinist decrees. The Veps intelligentsia were purged and the textbooks were confiscated (Strogaľnikova 1990).

were conducted in the 1920s and early 1930s, although the repression described above significantly reduced the number of teachers for the Mordvin-language schools (Kreindler 1985b; Lallukka 1990). The availability within the Mordvin Republic of the Mordvin languages as languages of instruction went from a high point of availability in grades 1-10 at the end of the 1930s, to availability in grades 1-7 up to 1959, to a low point of availability in 1972, when instruction using Mordvin was only provided up to grade 3. It was still possible to study Mordvin as a subject up to grade 9, however. Now some schools are again offering instruction in Mordvin through grade 4 (Lallukka 1990; this volume). But the figures are not encouraging. During the 1980s the number of children learning a Mordvin language as a subject fell from 80,000 to 24,000, while the number of children studying in Erzia or Moksha dwindled from 12,000 in 1980/81 to 4,000 in 1988/89 (Lallukka, this volume).

Language education and language maintenance among the Mordvins have been relatively poor in urban areas and quite poor in the Mordvin enclaves outside the Mordvin Republic. Up until recently, Erzia and Moksha were not used as languages of instruction nor even taught as subjects in elementary and secondary schools in Saransk, the capital of the republic (Rogers 1991, Lallukka, this volume). Lallukka also reports that there may have been no native-language schools at all outside of the Mordvin Republic in the 1980s.

A distinct danger is that Mordvins may have a lower view of the prestige of their language(s) than some other speakers of national languages in Russia, especially those living outside of their titular republics. For example, a report from the Orenburg Region (in southern Russia near the Ural Mountains and just north of Kazakhstan) on developing local languages reports that "in population centers where there are compact communities of Kazakhs, Ukrainians, Mordovians, and Chuvashes, their native languages and literature are not being studied, although the majority of the people polled (74.8 percent), *with the exception of the Mordovian population*, favor their study" [emphasis added] (Bureau of the Orenburg Regional Committee of the CPSU 1990). We cannot tell from this statement just how many, or how few, Mordvins favored the study of a Mordvin language, but it is indicative of other areas that the Mordvins have a noticeably lower commitment to use of their language than certain other groups. Clearly,

all these facts present a poor outlook for the use of the Mordvin languages across a broad functional spectrum.⁴

5.1. Attempted renaissance for Mordvin language and culture. On August 3, 1990, the First All-Union Congress of the Society for National Rebirth began in Saransk, capital of the Mordvin Republic. The Society itself is named *Mastorava*, 'Motherland.' This congress brought together Erzia and Moksha representatives from throughout the Soviet Union. A Congress Resolution was prepared on August 4, 1990. This Resolution, along with the *Mastorava* program and regulations, was published in the July-August 1990 issue of the literary and political journal *Mokša*. Since the Resolution, in particular, gives virtually the only glimpse of the aims of this movement that we have, it is worthwhile to consider the proposals put forth in it.

5.2. Principles, plans, and goals of the *Mastorava* movement. According to the Resolution of the *Mastorava* congress, "the preservation and development of the Erzia and Moksha people [is] the main direction in the activities of '*Mastorava*'" (Rezoljucija 1990:10). The Resolution contains several plans for improving Mordvin political, cultural, and linguistic status. We are mainly concerned here with those recommendations which can be broadly construed as efforts at improvements in language status. These can be grouped roughly into four categories: improvements in language education; broadening of the scope of language use; creating stronger roots for language use through increased cultural education; and increasing the agricultural-industrial base in the Mordvin Republic.

5.2.1. Improvements in language education.⁵

•Introduce instruction in Erzia and Moksha in preschools and elementary schools and teach Russian as a subject. Promote bilingual education at the secondary school level and above.

⁴At the risk of stating the obvious, I will mention that upper-level schooling is not an essential prerequisite to language maintenance, since home or "kitchen" varieties of language can go on almost indefinitely. This fact seems to be rarely acknowledged in most studies of native-language educational opportunities. However, the type of language that is used in domestic settings, of course, is not adequate if a group wishes to conduct its political and economic life in the ethnic language.

⁵Since the Resolution was prepared in 1990, it refers to the MASSR (Mordvin Autonomous Soviet Socialist Republic) and certain other institutions that may no longer exist or be designated as they were at that time.

5.2.2. Broadening of language scope.

•Increased publication:—Ensure monthly publication of *Mokša* (a journal published in Moksha) and *Sjatko* (a journal published in Erzia) — Increase the size of the children's publications *Pioneren' vaig'el'* ('Voice of the Pioneer') and *Jakst'er' Täštenä* ('Little Red Star'). The Resolution asked that the titles of these be changed, a reaction against "Sovietization" of cultural content (cf. Kirkwood 1989). This has undoubtedly already taken place independently of action by *Mastorava*. — Encourage publication of republic and raion newspapers in three languages (i.e., Moksha and Erzia in addition to Russian).

•Instatement or reinstatement of native toponyms and ethnonyms: — Change official designation of the Mordvin (*Mordovskij*) nationalities to "Erzia and Moksha nationality". — Rename the Mordvin ASSR the "Erzia and Moksha ASSR" — "Reestablish destroyed toponymy" in the MASSR and other Moksha and Erzia regions; i.e., reinstate the Mordvin names for localities and geographical features — Register the "ancient" (Mordvin) names of localities and rivers in official documents

•Establish "parity bilingualism" in the MASSR (i.e., reduce the role of Russian in public life and put the Mordvin languages on an equal footing with it).

5.2.3. Increase in cultural education.

•Open a department of Finno-Ugric studies and a department of national culture at N.P. Ogarev Mordvin State University in Saransk

•Include the study of Erzia and Moksha folklore, history, philosophy, etc., in public schools and institutions of higher education.

5.2.4. Increase of the agricultural-industrial base in the Mordvin Republic.

A fourth area of action which I am construing as relevant to language policy may not at first glance seem to be pertinent to this problem, and in fact was probably not proposed by the Resolution's authors for this purpose. That is the enlargement of the agricultural-industrial base in Moksha and Erzia villages. To understand the significance of such action for language revival, it is necessary to invoke Andrzej Porebski's "blueprint" for language revivals (Porebski 1987). He notes that ethnic groups can reverse their peripheral position within the structure of a state by attempting to restore regional methods of production, using their own ethnic language. Providing further work opportunities in the

ethnic territory reduces one of the forces which compels ethnic group members to leave their homeland. At the same time, maintaining or restoring traditional technologies generally provides an environment in which existing native terminology can be used. Porebski's claim that industries utilizing the native language can either reduce the appeal of in-migration for new dominant-language speakers, or force those in-migrants to adopt the regional language, would not appear to be valid for the Mordvin case, since most Russian in-migrants tend to come to urban areas to work in higher-technology fields, which can't be expected to convert to Mordvin-language operations.

It is impossible to know how much research the *Mastorava* organizers did into the history of other language maintenance and revival efforts, but it is interesting to note that the *Mastorava* plan follows quite closely the findings of Porebski based on the experience of Basque, Catalan, and Romansh revival movements. He states that ethnic groups can try to improve their position vis-à-vis a dominant culture and language by (a) pressing demands for regional administration, (b) establishing ethnic political parties, (c) attempting to restore regional methods of production, (d) demanding schooling in their ethnic language, and (e) giving legal status to the language of the ethnic group equal to that of the dominant group. All of these points are covered in the *Mastorava* Resolution.

6. Accomplishments and actions since 1990. Regrettably, there has been no accessible trace of this movement, or of other progress in the area of an increase in Mordvin language status in the year and a half since the convention. While earlier on it was possible to suppose that the lack of information to report was due to the short amount of time that had passed since the congress, the greater period of time that has since elapsed, along with the economic chaos now prevailing in the Russian Federation, makes it unlikely that the movement will continue. Various inquiries to Mordvin scholars and other Finno-Ugrists in Russia have consistently failed to produce information about the activities of *Mastorava*. Recently Yevgeni Khelimskii, a prominent Uralic scholar in Russia, noted that there has been no visible progress with the *Mastorava* movement, and that this is probably due to a loss of interest in the plan. "Maybe it means that the '*Mastorava*' is just ... one of innumerable cultural [and] social initiatives and societies, which were growing lately in this country, just like mushrooms after rain [i.e., the

initiation of glasnost—JMHJ” (Khelimskii, personal communication). The one positive piece of news is that the Mordvin Republic was declared officially trilingual—using Moksha, Erzia, and Russian—after heated debate by the republic’s parliament on December 11, 1990 (*Pravda*, Dec. 11, 1990).

In the absence of more details about implementation of these plans, I would like to discuss issues of their motivation and chances for success.

7. The role of language revivals at the end of the Soviet era. Many autonomous republics declared sovereignty in the months before the August 1991 coup, and a change in the status of the local language went along with it. Among the titular Finno-Ugric republics, the parliament of the Udmurt (Autonomous Socialist) Republic declared sovereignty and named Udmurt as co-state language in 1990 (*Izvestija*, Sept. 22, 1990). The Supreme Soviet of the Komi (Autonomous Socialist) Republic voted to drop “autonomous” from the name of the republic, and submitted a proposal to make Komi a co-state language (*Izvestija*, Aug. 31, 1990).

There were many reasons to want political and economic autonomy in the USSR, not the least of which was a desire not to be dragged down with the obviously sinking Soviet ship of state; looked at more positively, what local entities desired primarily was freedom from central political and economic control (and only in some cases from local Russian-majority control), and the opportunity to benefit economically from local natural resources and industry. With such underlying motives, language pride may have been window dressing for some politicians and demagogues, an indication of “ethnic sincerity,” which was seen as more morally compelling than “selfish” economic considerations. An argument in favor of this view is the fact that other than in the former Estonian SSR and the Komi-Permiak Autonomous Okrug, none of the Uralic peoples forms a majority in its own administrative area. Clearly, therefore, the presumably Russian-heavy legislative bodies in these republics did not approve the declarations of sovereignty with the expectation that the Russian-speaking majority would learn the local language, but because there were enough other expected benefits to make it worthwhile for them to compromise and support local-language sentiments by the ethnic minority.

The authors of the *Mastorava* materials were undoubtedly sincere in their desire for an increased role for their native languages. Still, political

aims are promoted just as strongly in these same materials. It may be that we should regard the "language revivals" of 1990-91 as isolated products of this period of flux, and not have great expectations for any of them.

8. Outlook for the future of the Mordvin languages. The Mordvin situation is paradoxical. Although it may be stated that Erzias and Mokshas show a degree of disinterest in their languages in survey responses (cf. the Orenburg report above and data cited in Rogers 1991), there is nevertheless very high Mordvin language use and retention, especially within the Mordvin Republic. In the 1989 census, about 88% of Mordvins within the Mordvin Republic reported speaking their ethnic language (i.e., either Erzia or Moksha, not distinguished in the census) as a native language. Sixty-seven percent elsewhere in the USSR reported Mordvin as a native language (Lallukka 1990).

As Kreindler (1985a, 1985b) points out, authors have been predicting the doom of the Mordvins for some 200 years, but they seem to have missed something essential about Mordvin solidarity and ethnic identification. As Kreindler states, "Perhaps the mere fact that 862,847 people who derive no advantage from their Mordvinian self-identification [i.e., Mordvins outside the Mordvin Republic—JMH] and can physically pass as Russians, have still declared themselves as Mordvinians is in itself not without significance" (Kreindler 1985b:255). So the Mordvins persevere in their ethnic awareness. This does not guarantee continued use of Moksha and Erzia, however; Keskitalo (1981) points out that knowledge of Sámi (Lapp) is low in the reindeer-herding areas of Scandinavia although "Sámi *identity* is not stigmatized or kept hidden" (p. 157) [emphasis in original].

Although the old-world lay view of ethnicity counts language use as an indispensable criterion for membership in an ethnic group (Harlig 1990a), it is nevertheless true that an individual may consider himself as belonging to a given ethnic group without having any appreciable knowledge of the original language spoken by that group. The Mordvin case itself demonstrates this: only about 774,000 (~67%) of the 1.15 million Mordvins in the 1989 census claimed Mordvin as their native language (Künnap 1991). This means that one-third of the Mordvins are content to think of themselves as ethnically Mordvin with little or no knowledge of Moksha or Erzia.⁶ So the Mordvins could continue

⁶However, as Bill Fierman (personal communication) points out, most Soviet

as a self-defined ethnic group, maintaining certain customs, for example, without continuing to use Mordvin or Erzia in their daily lives. While this outcome may be abhorrent to current protectors of ethnic identity in this region, it is an increasingly familiar and "unmarked" situation in industrialized countries of the West, as in the case of most third-generation offspring from immigrant communities in the United States.

My guess is that Mordvin will continue to be in use for some time to come, but its use will not be due to the top-down planning of education officials and the literary elite. Increased interest on the part of the intelligentsia will not in itself lead to the preservation of Erzia and Moksha as living languages with a wide functional range. There will be limited success in expanding the prestige and range of social functions carried by Mordvin.

It may turn out that the *Mastorava* movement—or at least, the memory of it—and serious fears about the future of Mordvin may bring about efforts at consolidating the Moksha and Erzia dialects and their speakers. Kreindler notes that "one possible positive factor in the scattering of the Mordvins is that outside their original territory the Moksha and Erzia mix more freely with each other and are slowly consolidating in culture and language" (Kreindler 1985b:255). Language standardization generally refers to the creation of a supra-dialectal variant of a particular single language. The Erzias and Mokshas must decide if theirs is one language, or that Erzia and Moksha are, and should remain, distinct. If the Erzias and Mokshas see their fate as a common one, then they will have the best hope of achieving their goals by working on a unified "Mordvin" language. Consolidation would increase the political clout of the Mordvins as a whole, reduce by half the costs and effort required to produce native-language educational and administrative materials, and increase the audience for books, newspapers, and theatrical productions.⁷

citizens declared themselves to be the nationality indicated in their passports, which in turn was generally derived from the parents' nationality. Thus, there may be many fewer individuals who "feel" themselves to be Mordvin.

⁷The evidence of *Mastorava* points toward a decision to remain separate, at least in name. It appears that the motivated intellectuals among the Mordvins want to do away with the designation *Mordvin* altogether. As noted above, the Resolution calls for renaming the titular republic the "Erzia and Moksha Republic," and renaming the nationality the "Erzia and Moksha nationality."

9. Outlook for the maintenance of the non-Slavic languages of Russia. The non-Russian peoples of the Russian Federation, including the Mordvins, face a future which is unpredictable. The break-up of the Soviet Union, the dissolution of Communist Party control, and the change to a market economy can as easily be detrimental to minorities in Russia as helpful. We should remember that there was a certain power to being a "Soviet nationality." The smaller cultures and their languages, while all subordinate to Russian, were on a relatively equal footing with regard to one another. Provided a given group could survive the periods of destructive policy emanating from the center, a constant, albeit not generous, flow of funding was available for research and publication on language and culture. The current chaotic situation of a shattered economy and interethnic tension means that small languages must compete with each other for resources as well as with Russian for prestige.

It is easy to arrive at a very pessimistic vision of the prospects for maintenance of ethnic identity, particularly language retention, in the coming years. It is ironic that processes of democratization need not benefit the Mordvins or other groups. The change to a market economy has already started a (hopefully, short-term) reversal of fortune for what language reform and revival plans were afoot before the August coup, since the economic chaos which has now ensued leaves less funding than ever available for broadcasting, publication, and education in the minority languages.

The situation of these languages may not improve even when (or if) a stable market economy is in place. On one hand, if publishing, in particular, becomes fully demand- rather than command-based (a process which had already started under perestroika), the small numbers of interested and sufficiently literate readers for the various languages may make native language printing houses (e.g., the Mordvin and Mari Book Publishing Houses) unprofitable and hence unviable.

On the other hand, an upswing in the economy does not guarantee improvements, either. The Russian language up to now has been not so much the bearer of greater opportunity for minority speakers as a means simply of having life a little less bad; thus it has been an option that some speakers—particularly those in the rural-agricultural and other less-skilled segments of society—have been willing to pass up. But a Russian language "bearing gifts," i.e., bringing the promise of

an automobile, private home, foreign travel opportunities, and other perks of capitalism (whether attainable in reality or not) to minority speakers will be much harder to resist.⁸ Beyond this, English, in connection with trade in new goods and an increasing spirit of entrepreneurship, could begin to infiltrate even more significantly than it already has: it is an ironic fact that, in effect, Russian has formed a protective barrier for the languages of the former USSR against the otherwise inexorable encroachment of English world-wide. Finally, in spite of terrible ecological damage done to many Uralic homelands (particularly the gas- and oil-rich Yamalo-Nenets Okrug in western Siberia [Tropkin 1990, Batashev 1990]), the reality is that massive inefficiency in the Soviet economic system has, up to now, protected many areas with natural resources from even more serious exploitation. A driving, efficient Russian economy could devour forest, steppe, and agricultural land at a prodigious rate. Loss of these lands means a further loss of traditional means of subsistence, leading in turn to further urban migration. Loss of traditional technology means another area where the native language loses a function and a foothold.⁹

10. Sociolinguistics, former Soviet nationalities and future studies. Several factors have made language use among former Soviet peoples inaccessible to direct study and observation. First, few Western scholars have been permitted to travel to the areas where most speakers of languages other than Russian live. Second, the methodology of Western sociolinguistics, which includes analysis of face-to-face interaction and other microsociolinguistic phenomena, did not find a favorable reception in the Soviet Union; in fact, the prevailing paradigm was outright hostile to it (cf. Švejcer 1986, Harlig 1990b). Quite naturally, therefore, sociolinguistic studies of Soviet nationalities have concentrated on quantifiable data available from census reports and book publishing data (e.g., *Knižnaja letopis'*). This macrosociolinguistic approach to the investigation of language use can be thought of as a

⁸Naturally, the idea of a thriving Russian economy requires a good deal of imagination given the current crisis.

⁹My experience in the Mari Republic suggests that the collective farm system actually aids in the preservation of language and culture, through folk dramatic and musical ensembles sponsored by the collectives. But giant mechanized farms, to which large areas of Russian agricultural land are well-suited, would destroy even this degree of social cohesion.

historical-institutional method. However, this approach ignores a fundamental fact about language choice: the decision by bilingual speakers to use the dominant vs. the local ethnic language is made at the level of conversational dyads and small groups in particular speech situations. The overall increase or decrease in the use of an ethnic language is the result of individual decisions made in innumerable encounters. The use or non-use of the ethnic language is generally divided by functions, including political administration, performance of workplace duties, broadcasting, publishing, discussing domestic affairs, and conducting religious services. Atrophy of language use takes place at the functional level before it occurs as attrition of language use or ability in individual speakers. Most existing studies fail to make this distinction, however, looking only at large measurable societal trends.

Whatever the level being studied, it is clear that self-reports of proportion and location of language use, such as those available in the *Vestnik Statistiki*, are highly inaccurate. In the first place, individuals have difficulty keeping track of language choice; as with many human behaviors, our intuitive perception may vary widely from what trained observers would find. Second, responses are susceptible to all kinds of political and social effects: the most obvious skewing occurs when ethnic minorities minimize their reports of native language use either for fear of marking themselves for possible persecution from the state, or because of low self-esteem caused by belief in the low social standing of their native language compared to the dominant language. But skewing can also take place in the opposite direction, when ethnic revivals cause increases in ethnic pride and awareness. Then respondents may exaggerate (even if unconsciously) the amount or location of their native language use.

Due to these shortcomings, historical-institutional statistics and self-report data must eventually be largely replaced by scientifically constructed surveys and studies of face-to-face interactions. Within a relatively short time, Western scholars, along with Russian and native researchers no longer bogged down by Marxist and Soviet dogma will be able to begin the microanalysis necessary to determine the real factors involved in language choice. We should be ready to take advantage of the opportunity.

REFERENCES

- Batashev, A. 1990. A fragile corner of Siberia. *Soviet Union*, no. 10 (487):18-21.
- Bureau of the Orenburg Regional Committee of the CPSU. 1990. On the work of the Asekeyevo District party committee to improve inter-ethnic relations and develop local languages and cultures. Decision of the Bureau, September 16, 1989. *Current Politics of the Soviet Union* 1.2.164-166.
- Frank, Allen. 1992. Translation of "Resolution of the First All-Union Congress of the Mordvin *Mastorava* Society for National Rebirth." *Ural-Altaische Jahrbücher* 64.153-155.
- Hajdú Péter. 1980. *Uráli nyelvrokonaink*. [Our linguistic relatives, the Uralic peoples.] Budapest: Tankönyvkiadó.
- Harlig, Jeffrey. 1990a. Theoretical problems in the description of American Hungarian. Paper presented at the conference "National Identity and Culture: Hungarians in North America," Indiana University, Bloomington, April 2, 1990.
- _____. 1990b. Prospects for change in East European social science. *Proceedings of the fifteenth European studies conference*, ed. by Karl Odwarka, 101-108. Omaha: University of Nebraska.
- Izvestija*, Aug 31, 1990. Reported in *Current Digest of the Soviet Press*, vol. 42, no. 35, 1990.
- Izvestija*, Sept. 22, 1990. Summary in *Current Digest of the Soviet Press*, vol. 42, no. 38, 1990.
- Kirkwood, Michael. 1989. Language planning: Some methodological preliminaries. *Language planning in the Soviet Union*, ed. by Michael Kirkwood, 1-22. London: Macmillan.
- Kreindler, Isabelle T. 1985a. The Mordvinians: A doomed Soviet nationality? *Cahiers du monde russe et soviétique* 26.1.43-62.
- _____. 1985b. The Mordvinian languages: A survival saga. *Sociolinguistic perspectives on Soviet national languages*, ed. by Isabelle T. Kreindler, 237-264. Berlin: Mouton de Gruyter.
- Künnap, Ago. 1991. Uurali keelte valdajaskond Nõukogude Liidu poolt kontrollitaval territooriumil 1989. [Command of the Uralic languages in Soviet-controlled territories, 1989.] *Keel ja Kirjandus* 34.7.416-419.
- Lallukka, Seppo. 1990. *The East Finnic minorities in the Soviet Union: An appraisal of the erosive trends*. [Annales Academiae Scientiarum Fennicae, Series B, 252.] Helsinki: Suomalainen Tiedeakatemia.
- _____. 1993. The status of the Volga-Finnic languages in comprehensive school. *This volume*, 185-210.
- Popov, A.A. 1990. How did "blank spots" occur in the history of national regions? Roundtable on the nationality question. *Soviet Studies in History* 29.1.85-87.
- Porebski, Andrzej. 1987. Language and ethnic groups. *Language and nation*, ed. by Władysław Miodunka, 33-45. Warsaw/Cracow: Jagiellonian University.
- Pravda*, Dec. 11, 1990. Reported in *Current Digest of the Soviet Press*, vol. 42,

no. 50, 1990.

Rezolucija 1-go vsesojuznogo s"ezda obščestva nacional'nogo vrozroždenija "Mastorava." [Resolution of the First All-Union Congress of the Society for National Rebirth, "Mastorava".] 1990. *Mokša*, Jul.-Aug., 9-11.

Rogers, Greg S. 1991. The Mordvin language revival: Does it have limitations? Ms., Department of Uralic and Altaic Studies, Indiana University.

Sanukov, Ksenofont. 1991. Tragedija. [Tragedy.] *Marijskaja Pravda*, January 15, 2,4; January 16, 2, 4; January 18, 3-4; January 22, 2, 4.

Strogal'nikova, Z. I. 1990. The Veps: History and problems of ethnic development. Roundtable on the nationality question. *Soviet Studies in History* 29.1.73-76.

Tropkin, Alexander. 1990. Yamal: What price progress? *Soviet Life*, December, pp. 16-31.

Švejcer, A. D. 1986. *Contemporary sociolinguistics: Theory, problems, methods*. Amsterdam/Philadelphia: John Benjamins.

Khelimskii, Yevgeni. Private letter, August 7, 1991.



ERGATIVE-TO-ACCUSATIVE SHIFT IN AGREEMENT: TABASSARAN¹

Alice Carmichael Harris
Vanderbilt University

1. Introduction. The alignment of agreement, like that of cases, may be of several types. Three important types are summarized in Table 1.

| | Dir Obj | Subject Intr | | Subj Tr |
|------------|---------|--------------|--------|---------|
| | | Inactive | Active | |
| Ergative | A | | A | B |
| Active | A | A | B | B |
| Accusative | A | | B | B |

Table 1. Alignment Types (after Sapir 1917).

The term alignment is used here as a neutral designation of these types. Changes in the alignment of case marking have been studied in languages of several families; examples include Anderson (1977) on certain Indo-Iranian languages, Chung (1978) on some languages of the Polynesian family, Klocke (1978) on some Pama Nyungan languages, and Harris (1985) on the Kartvelian languages. On the basis of these and other studies, we may say that the means by which languages change case alignment type are now reasonably well understood; these shifts typically retain the same core set of case markers, and the mechanisms are typically reanalysis of a syntactic construction or extension of cases. Changes in the alignment of verb agreement, on the other hand, have been little studied, except where they occur as concomitants of case changes of the types above. In comparison with change in the alignment of case marking, independent shifts in the alignment of agreement are poorly understood.

A few of the Daghestan languages (North East Caucasian family)

¹This material is based upon work supported by the National Science Foundation under Grants Number BNS-8217355 and BNS-8419143. Earlier versions of parts of this paper have been presented to the Annual Meeting of the Linguistic Society of America, the Washington Linguistics Club, and the Symposium on Inflectional Morphology and Syntax held at the University of North Carolina at Chapel Hill. I am grateful to these audiences for helpful comments.

have undergone independent ergative-to-accusative shifts in agreement, and in this paper the diachronic processes involved in such a shift in a language of the Lezgian branch of this family are studied, with special attention to mechanisms of the alignment change. It is argued here that the inherited set of markers was supplemented, then replaced, by a new set (derived from personal pronouns) and that the mechanism involved was reanalysis of a clitic copy pronoun originally marking the topic. As a result of this analysis, the process involved in these shifts of agreement alignment is seen as very different from those involved in analogous instances of shift of case alignment.

Scholars generally agree in reconstructing for Proto-Daghestanian a system of ergative case marking and **ergative agreement**, coding the **grammatical gender-class** (defined below) and number of the **absolutive -nominal**.² In general in languages of this family, gender-class agreement is realized through **prefixes**, and many believe that this represents the original situation. A few of the Daghestan languages have added an agreement rule which has (nominative-) **accusative alignment**. The accusative agreement rules generally involve **suffixes** which code the **person** and **number** of the **subject**. Because agreement with accusative alignment began to coexist with or replaced agreement of ergative alignment, these changes may be considered examples of ergative-to-accusative shift.

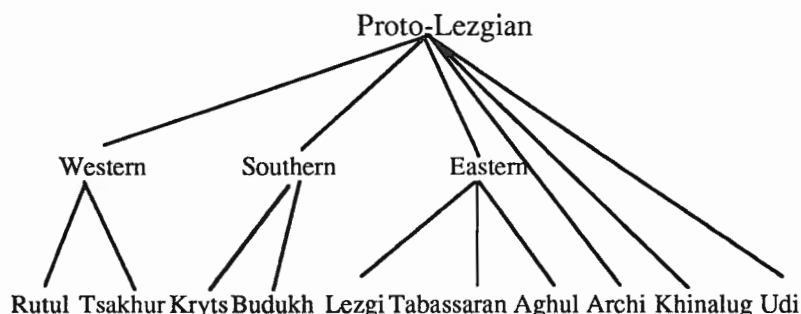
The term **ergative agreement**, as used here, refers to the alignment of the rule, in the sense of Table 1. It does *not* necessarily refer to agreement conditioned by the ergative-marked nominal, and in fact in each instance discussed here it is the absolutive nominal that triggers the rule. Thus ergative agreement in Daghestan languages is agreement conditioned by direct objects and by subjects of intransitives, the "A" relations of the ergative in Table 1. Similarly, **accusative agreement** does not necessarily refer to agreement conditioned by the direct object; in fact in each instance discussed here the markers are conditioned by

²Some linguists who have made this reconstruction have done so on the basis of their assumption, considered invalid by most western Caucasologists, that the North East Caucasian family is related to the North West Caucasian and Kartvelian (South Caucasian) families (eg., Andžulaze 1968, Čikobava 1962). Nevertheless, there is a firm basis for reconstructing for North East Caucasian the system described here without assuming genetic affiliations that Čikobava and his followers support.

the subject. The innovative agreement in Tabassaran is conditioned by transitive and intransitive subjects, the "B" relations of accusative in Table 1. Thus, both terms refer to the alignment of the rule, in the sense of Table 1, not to the nominal that conditions the rule.

The North East Caucasian family comprises the Lezgian, Avaro-Andi-Dido, Lak-Dargwa, and perhaps the Nakh branches, though some linguistics consider the last of these to constitute a distinct family.

The Lezgian branch is composed of ten languages: Lezgi, Tabassaran, Aghul, Rutul, Tsakhur, Archi, Kryts, Budukh, Khinalug, and Udi. The tree below shows relationships among them, omitting some details.



This paper focuses on changes in agreement in Tabassaran. In interpreting the change in Tabassaran, the paper also draws on the Lezgi language and on Aghul, both closely related to Tabassaran.

Section 2 of this paper presents a description of agreement in Lezgian languages and specifically in Tabassaran. The third section continues this description, at the same time presenting previous accounts of the shift in agreement. Section 4 proposes a unified account of the shift in agreement alignment.

2. Agreement in Lezgian Languages. Grammatical gender-class (henceforth also **class**) is used here to refer to a system of classifying or typing nouns and coding these types grammatically. Specialists agree in reconstructing four gender-classes for the singular in Proto-Lezgian; with approximate characterization of their semantics, these are I. masculine, II. feminine, III. other animate, IV. inanimate (Andyulage 1968: 35-39; Xajdakov 1980: 186ff.; Kaxaze 1983;

Alekseev 1985:89-91). In Proto-Lezgian and in some modern languages of the Lezgian group, class is used (a) to mark agreement of the verb with the absolutive nominal, (b) to mark concord of modifiers with head nouns, and (c) to a limited extent in word formation. It should be added that there was originally no person agreement (agreement according to first, second, and third persons) in Proto-Lezgian or in North East Caucasian; it is an open question whether number agreement was in the proto-language at any stage.

Among languages of the Lezgian group, both Tabassaran and Udi have developed extensive systems of agreement with accusative alignment.³ As shown in the tree above, these languages are not closely related, and agreement appears to have developed entirely independently in the two, although there are some similarities in the systems.

In the northern dialect of Tabassaran, two systems of agreement coexist. The inherited system involves the coding of the gender-class and number of the absolutive nominal by means of prefixes on the verb. The original set of four classes has been simplified to two in Tabassaran: human vs. non-human. Markers of the human class are *d*, *r*, *y*, and \emptyset ; markers of non-human include *b*, *w*, *f*. This agreement system is illustrated in (1) and (2).⁴

(1) (a) *izu d-isnuza baj*.

I.ERG HUM-catch boy.ABSL

I caught a boy.'

(b) *izu b-isnuza žaq'a*. (Magometov 1965:198)

I.ERG NON-catch bird.ABSL

³Tsakhur (Ĵeiranišvili 1953:458), Khinalug (Dešeriev 1959:45), and Kryts (Topuria 1960:439) appear also to show rudimentary differentiation of verb forms according to the person (and number) of the subject.

⁴Abbreviations used in the glossing of examples are the following: Erg = ergative case, Absl = absolutive case (called nominative in some sources), Dat = dative case, Loc = locative case; Hum = human class, Non = non-human class; PV = preverb; 1.Sg = first person singular subject marker, 2.Sg = second person singular subject marker, 3.Sg = third person singular subject marker; AOR = aorist. Although there is syncretism in the absolutive and ergative forms of the first and second person independent pronouns in Tabassaran, in examples pronouns are glossed with one case, according to their function. Although the third person pronoun does not distinguish gender, 'he' is used to gloss these pronouns throughout.

'I caught a bird.'

(2) (a) *li-Ø-Xnuv*.

'He (human) worked.'

(b) *li-v-Xnuv*. (Kibrik and Seleznev 1982:18)

'He (animal) worked.'

In the transitive (1) sentences, the direct objects condition agreement according to their class—human or non-human. In the intransitive (2) sentences, it is the subject that conditions class prefixes. This kind of agreement, however, is not productive (Topuria 1962:63, citing Žirkov). The innovative agreement system is independent of the inherited system. The new system codes the person and number of the subject by means of suffixes on the verb; this is illustrated for transitive sentences in (3) and for intransitive in (4).

(3) (a) *izu b-isnu-za žaq'a*.

I.ERG NON-catch-1.SG bird.ABSL

'I caught a bird.'

(b) *iwu b-isnu-wa žaq'a*.

you.ERG NON-catch-2.SG bird.ABSL

'You (SG) caught a bird.'

(c) *duyu b-isnuw-Ø žaq'a*. (Magometov 1965:198)

he.ERG NON-catch-Ø bird.ABSL

'He caught a bird.'

(4) (a) *izu ti-r-xnu-za*.

I.ABSL PV-HUM-fly-1.SG

'I flew.'

(b) *iwu ti-r-xnu-wa*.

you.ABSL PV-HUM-fly-2.SG

'You flew.'

(c) *dumu ti-r-xnuw-Ø*. (Magometov 1965:197)

he.ABSL PV-HUM-fly-HE

'He flew.'

Note that in the intransitive, (4), the subject is coded for class and number by the prefix, here *-r-* of the human class, and for person and

number by the suffix. These examples illustrate also that the third person subject has a zero-marker, the final *w* in (3c) and (4c) being part of the verb stem.

The southern dialect of Tabassaran has lost the category of gender-class and with it the class agreement illustrated in (3-4); this dialect has only person-number agreement with accusative alignment, the innovative system.

As shown in greater detail below, the agreement suffixes have developed from the system of personal pronouns.

3. Previous Accounts of the Shift in Agreement. Consideration of the languages of the Lezgian group reveal certain correlations between accusative agreement and other grammatical phenomena. Some authors state or imply that one or another of these is crucially involved in the ergative-to-accusative shift in some way. Because of their possible explanatory value, it is important to examine these correlations.

3.1. Position of Agreement Marker. In general, there is a correlation in this language group and others of the North East Caucasus between position and alignment, such that ergatively aligned agreement is marked with prefixes and accusative agreement with suffixes. As shown above in section 2, this is the situation found in Tabassaran. However, the correlation is not absolute; in Udi markers of **accusative** agreement may occur as (non-initial) **prefixes**, as illustrated in (5).⁵

- (5) *ba-ne-k-i*
 *be*¹-3.SG-*be*²-Aor
 ‘he/she/it was’

Historically, *ba-* in (5) is a fossilized class marker (Ėiranišvili 1956), and both it and the agreement marker, *ne-*, precede the original root, *k*. The correlation between position of an affix and the alignment of the rule assigning it is thus not absolute in the Lezgian languages. Perhaps more importantly, there seems to be no sound reason why position of the marker should be relevant to the development of accusative agreement. However, it does appear that in the Lezgian languages suffixal agreement of any type is a relatively recent innovation.

⁵It is assumed for present purposes that agreement markers in Udi are true prefixes, rather than proclitics.

3.2. Characteristic Coded. In general, in this language group and others of the North East Caucasus there is a correlation between alignment of agreement and the characteristic(s) coded, such that ergative agreement codes for noun class and number, while accusative agreement codes for person and number. As shown in section 2, this is characteristic of Tabassaran. However, again the correlation is not complete. Archi has agreement that has ergative alignment, coding the person (third) and number (plural) of the absolutive nominal, though this phenomenon is limited to this person and number (Kaxaḡe 1979: 218-20). Thus, there is not an absolute correlation between ergative agreement and coding for class and number alone. Nevertheless, in Archi, only one distinction is made with respect to person: third plural is distinguished from all others. It does appear that coding for multiple distinctions (first vs. second vs. third) by a rule of accusative alignment is accomplished in languages of this group only by means of affixes that originated as pronouns. Nevertheless, no hypothesis has been advanced that would specify the exact nature of the relationship between personal pronouns and agreement of accusative alignment.

3.3. Polypersonalism. With respect to the North West Caucasian and Kartvelian languages, Rogava (1962:50) has advanced the hypothesis that the development of polypersonalism, that is, agreement with more than one nominal, is related to what he views as a change from class to person agreement in those families. Yet when we consider languages of the North East Caucasian family, we see immediately that polypersonalism cannot be a necessary correlate of the class-to-person shift, inasmuch as Udi, which has fully developed person agreement, gives no indication that it has ever used that system to mark agreement with more than one nominal (see Ĵeiranišvili 1956, 1971; Pančviḡe 1974, Schulze 1982).

3.4. Loss of Absolutive Agreement. Of the ten Lezgian languages, the following have lost the system of grammatical class: Lezgi, Aghul, Udi, and the southern dialect of Tabassaran. While the Lezgi language, Aghul, and Tabassaran form a subgroup within the Lezgian group, Udi is not closely related to them. In the northern dialect of Tabassaran the number of classes has been reduced from the original four to two, and the use of the markers is not truly productive. Of the remaining six languages, all but Archi have made rather extensive changes in the system. On the basis of these facts, it is said that the system of

grammatical class is weakening in Lezgian (for example, Andžulage 1968:32; Kaxaze 1983:320; Alekseev 1985:89).

The loss of the system of grammatical class entails the loss of the inherited rule of class agreement, which had ergative alignment. Although some of the languages which lost ergative agreement did not develop any other agreement system (Lezgi and Aghul), the loss of the inherited rule could be a factor in the development of a new agreement rule. Both of the languages that have developed extensive accusative agreement (Udi and Tabassaran) are among those that have lost or severely restricted the inherited agreement system.

3.5. General Accounts. Since changes in the agreement system from ergative to accusative, from class to person, or from prefixing to suffixing have occurred in some of the languages of each branch of the North East Caucasian (NEC) family, there may have been a tendency toward such change already present in Proto-NEC before splits into distinct groups and languages. Some linguists have attributed the changes in part to "the development of abstract thinking" or to the fact that these peoples have attained a "higher cultural level" (Čikobava 1937:103; Javaxišvili 1937: 257; Kaxaze 1983: 320).

A different approach is taken by Topuria (1962: 62-3), who has suggested that person agreement (of accusative alignment) in Tabassaran originated because of polysemy in forms with only class (ergative) agreement. For example, *b-is*, with *b-* marking the direct object as non-human, would once have meant all of the following: 'I take it (non-human),' 'we (inclusive or exclusive) take it,' 'you (singular or plural) take it,' and 'he/she/it/they take it.' Suffixation of forms derived from first and second person pronouns would disambiguate these, leaving the unsuffixed form to mean third person; eg., *b-is-za* 'I take it,' *b-is-va* 'you take it,' and *b-is* 'he/she/it takes it,' etc. While this is doubtless true, the same observation applies to any language in which the verb agrees with fewer than two arguments; yet agreement with only one nominal seems to be a stable situation in many languages. Like the hypotheses mentioned above, this one does not address the question of why the innovative system provides agreement with all subjects, rather than just with the ergative nominal. That is, Topuria's hypothesis does not explain why the innovative agreement has accusative alignment.

Kibrik (1979:74-75), like most others who have treated the shift in

agreement in languages of the Caucasus, focuses on the class-to-person aspect of the change. For example, he hypothesizes that the coding of person provides a way of overcoming the limitations imposed by the syncretism of ergative with absolutive in the forms of the first and second person personal pronouns. He suggests that the innovative agreement developed as a way of coding both agent and patient in transitive clauses. Still, he does not address the question of why this personal agreement should have accusative alignment; the addition of agreement with just the ergative nominal would likewise serve the purpose of registering both subject and object in transitive clauses, and such a system would preserve the ergative alignment of other rules of the language.

In part, the changes may be due to inter-influences among indigenous languages. These changes are sometimes attributed, at least in part, to the influence of contact languages of the Indo-European and Turkic families, where agreement is accusative in alignment. Although at least some degree of diffusion is a likely cause in this instance, it does not provide an internal mechanism for the changes. Since an agreement system was not simply borrowed wholesale from a neighboring language, diffusion provides an incomplete account. None of the hypotheses advanced to date addresses the question of the exact means by which accusative alignment could be introduced in languages that had ergative case marking and agreement.

In this section I have surveyed phenomena which, because of their apparent correlation with accusative agreement, have been assumed to be relevant to the development of this rule. Some correlations are spurious or incomplete, but some factors previously identified seem to play a genuine role in the change. I have described explanations previously advanced for this change. No account of which I am aware develops a clear hypothesis of why these correlations might exist or what bearing other grammatical phenomena have on the ergative-to-accusative shift. No account of which I am aware attempts to explain how or why the innovative agreement system developed accusative alignment, rather than renewing agreement of ergative alignment.

4. A Unified Account.

4.1. Developing the Hypothesis. In some dialects of Tabassaran, there is a process of optional clitic-copying in addition to the obligatory rule of subject agreement. In this process, a copy of a first or second

pronoun in almost any function may cliticize to the verb form. The functions in which the pronoun may occur in the sentence include direct object (semantic patient), indirect object (recipient), derived subject (stimulus) with affective verbs that undergo Inversion,⁶ and comitative, superessive, general locative, possessor in the subject, possessor in the recipient, possessor in the comitative, and possessor in the locative (Kibrik and Seleznev 1982:23, Magometov 1965:202, 208, 210, 211). The indirect object and one of the locatives are illustrated in (6) and (7).

- (6) *duyu izus b-ik'n-is k'aža*. (Magometov 1965: 208)
 he.ERG me.DAT NON-write-me.DAT letter.ABSL
 'He wrote a letter to me.'

- (7) *dumu quynu-zuq izuq*. (Magometov 1965: 211)
 he.ABSL believe-me.LOC me.LOC
 'He believed, trusted in me.'

As copies, these clitics reflect the case of the free pronoun. The optional clitic copy may occur with an obligatory subject agreement marker, as in (8).

- (8) *izu quynu-zu-č'uq ič'uq*. (Magometov 1965:211)
 I.ABSL believe-1.SG-2.PL.LOC you.PL.LOC
 'I believed, trusted you.'

I hypothesize that subject agreement in Tabassaran began in a similar way, optionally copying forms of the personal pronouns; these optional clitic copies of the pronouns were reanalyzed as obligatory markers, thus constituting agreement. I hypothesize further that historically the clitics were pronominal copies of the topic. While many kinds of nominals may be topicalized, it is well known that universally there is a high incidence of subject topicalization. We may assume that in Pre-Tabassaran, as in other languages, the subject was more often topicalized than other nominals, and that it therefore occurred as a

⁶In Inversion it is the initial subject that conditions obligatory person agreement in Tabassaran, as in Udi (see Harris 1984).

clitic copy more often than other nominals did.⁷ The optional clitic copies were eventually reinterpreted as obligatory markers of the subject.

It is only first and second person pronouns that may have a clitic copy in Tabassaran today, and the use of zero-agreement in the third person leads us to hypothesize that this was true also at earlier stages. This correlation between clitic copying and agreement provides support for the hypothesis that the former was the diachronic basis for the latter.

Topicalization thus appears to have been the mechanism for the change from ergative to accusative alignment; because topicalization applies especially to subjects, rather than absolutes, it provides a means for formally associating the subject of the transitive with the subject of the intransitive.⁸ It is likely that the reanalysis described above came in response to pressure from partial loss of the inherited class agreement. It may also be that influence of neighboring languages with subject agreement had a bearing on this process. Yet these two facts are not sufficient to explain the alignment shift in the absence of topicalization.

4.2. The Form of the Verbal Suffixes. In some southern dialects there is an important variant of the system described above. Here the subject of a transitive verb conditions obligatory agreement just as in the northern dialects, while the subject of an intransitive verb conditions an obligatory marker with a different form. The latter form is also used for the optional clitic copy of the direct object of a transitive verb (Kibrik and Seleznev 1982:25-26; Magometov 1965:198-201, 202). The two sets of markers found in the southern dialect are summarized in Table 2.

| | |
|---------------|-------------------|
| Subj of Trans | Subj of Intrans & |
|---------------|-------------------|

⁷Mallinson and Blake (1981:103-114) adduce evidence that absolutes are more highly topical in languages with ergatively aligned phenomena, but others disagree. For example, Du Bois (1985:351, also 1987:842) notes that agents and hence transitive subjects are frequently topics in Sacapultec; Plank (to appear, §2.2.3) states that an ACT_i (the active argument of a transitive verb) is "the topicworthier argument especially if verbs are imperfective or progressive and the INACT_i is less than fully affected..."

⁸The transition from clitic copy of a topic to marker of agreement is discussed in Givón (1976), but it is not noted there that this may result in an ergative-to-accusative shift.

| | | Dir Obj of Trans |
|---------------|------------------|------------------|
| 1st person SG | za | zu |
| 2nd person SG | wa | wu |
| 1st person PL | ča | x̃u ⁹ |
| 2nd person PL | č ^o a | č ^o u |

Table 2. Verbal Suffixes in the Southern Dialect of Tabassaran
(from Magometov 1965:198-199).

The different uses of these sets of suffixes are illustrated below, where they are glossed as ERG and ABSL, respectively.

- (9) uzu gak'wler urgura-za. (Magometov 1965:201)

I.ERG firewood.ABSL burn-1.SG.ERG

'I burn firewood.'

- (10) uzu urgura-zu. (Magometov 1965:201)

I.ABSL burn-1.SG.ABSL

'I burn, am on fire.'

- (11) uzu uwu bisura-za(-wu). (Magometov 1965:202)

I.ERG you.SG.ABSL catch-1.SG.ERG(-2.SG.ABSL)

'I catch you.'

Thus, in this dialect all subjects condition agreement, but the agreement markers used for subjects of transitives are different from those used for subjects of intransitives.

The northern dialect does not systematically differentiate the two sets of suffixes in examples like (9) and (10), where a single suffix occurs (Magometov 1965:201); however, this dialect does distinguish the two sets of forms in examples like (11), where two nominals are registered on the verb (see Magometov 1965:204-205). This is shown in (12), from a northern dialect.

- (12) izu iwu d-iržunu-za-wu. (Magometov 1965:204)

I.ERG you.ABSL HUM-bend-1.SG.ERG-2.SG.ABSL

⁹The independent personal pronoun distinguishes inclusive *ux̃u* from exclusive *uču* (Magometov 1965:169).

'I bent you.'

- (13) *duyu iwu d-iržunu-wa.* (Magometov 1965: 205)
 he.ERG you.ABSL HUM-bend-2.SG
 'He bent you.'

As (3-4) and (13) show, the northern dialect uses the *za* set (the first column in Table 2) for subjects of transitives, for subjects of intransitives, and for optional clitic-copies of direct objects when subject agreement is zero. According to data available in Magometov 1965 and Kibrik and Seleznev 1982, this dialect uses the *zu* set only as the optional clitic-copy of the direct object when obligatory subject agreement is a non-null morpheme.

Although it is not immediately clear which dialect is archaic in this respect, data from closely related languages provide an answer to this question. Tabassaran forms a subgroup with Aghul and Lezgi; neither Aghul nor Lezgi has any verb agreement, but the pronoun forms, listed in Table 3, are revealing.

| | Lezgi Independent Pronouns | Aghul Independent Pronouns | Southern Tabassaran Agreement Suffixes |
|-------------------|----------------------------------|--|--|
| ABSOLUTIVE | | | |
| 1st person SG | <i>zun</i> | <i>zun</i> | <i>zu</i> |
| 2nd person SG | <i>wun</i> | <i>wun</i> | <i>wu</i> |
| 1st person PL | <i>čun</i> | <i>čin</i> (EXCL) <i>šin/šin</i> (INCL) | <i>šu</i> |
| 2nd person PL | <i>k'yun</i> | <i>čun/kün/č'un</i> | <i>č'u</i> |
| ERGATIVE | | | |
| 1st person SG | <i>za</i> | <i>zaš/zun</i> | <i>za</i> |
| 2nd person SG | <i>na</i> | <i>waš/wun</i> | <i>wa</i> |
| 1st person PL | <i>čna</i> | <i>češ/čin</i> (EXCL) <i>šeš/šin/šin</i> (INCL) | <i>ča</i> |
| 2nd person PL | <i>k'yune</i> | <i>čweš/čun/kØn/č'un</i> | <i>č'a</i> |

Table 3. Comparison of Pronouns in Lezgi and Aghul with Verbal

Suffixes of Tabassaran. (Sources: Andžulage 1968:117; Magometov 1970:101-102; Magometov 1965:198-99)

The absolutive forms of the pronouns in Lezgi and Aghul and the verbal suffixes used for absolutive nominals (subjects of intransitives and direct objects of transitives) in the southern dialect of Tabassaran share the vocalism *u* in all forms except the first person plural pronouns in Aghul. The ergative forms of some of the pronouns in Lezgi and two of those in one dialect of Aghul (*zaš*, *waš*), as well as the verbal suffixes used for ergative nominals (subjects of transitives) in the southern dialect of Tabassaran (*za*, *wa*, *ča*, *č^oa*) share the vocalism *a*.¹⁰ These data provide a firm basis for reconstructing two case forms for first- and second-person pronouns in Proto-Eastern-Lezgian: an absolutive form with the vocalism *u* and an ergative with *a*.¹¹ Alekseev (1985:71) reconstructs distinct forms for absolutive and ergative in Proto-Lezgian; he does not reconstruct a consistent vocalism for absolutive at this level, though he notes that the vowel was important in distinguishing between the two cases (1985:72). He likewise draws attention to the secondary nature of the tendency toward syncretism of absolutive and ergative in the personal pronouns in several of the Lezgian languages (1985:73).¹² In the three languages of the Eastern

¹⁰The Lezgi second person singular pronoun is probably from an earlier **wan(a)*, and the *a* vocalism of the present day pronoun may or may not be a reflex of the vowel we are interested in here. The same may apply to the first person plural pronoun of Lezgi.

¹¹Kibrik (1979:74-75) considers additional types of intransitives and concludes that "the southern dialect of Tabassaran is tending to shift towards the active type," rather than being strictly ergative. This refers to the current distribution of the agreement markers with the *a* vs. *u* vocalism and is not at odds with the reconstruction of the *a/u* distinction in the present paper.

¹²Žirkov (1948:128) likewise suggests that Tabassaran independent pronouns represent a language-particular syncretism continuing the form that was earlier the absolutive, while the agreement markers in vocalism *a* are a continuation of the ergative. Andžulage (1968:118-119) cites evidence from Udi (where both ergative and absolutive have forms in *u*) and Tabassaran and argues that these data show, contrary to Žirkov's suggestion, that the ergative-absolutive distinction in the pronoun is an innovation specific to Lezgian. Andžulage, in 1968, may have been unaware of the *a* forms in closely-related Aghul (from Magometov 1970) and did not take the suffixes from southern dialects of Tabassaran into consideration. If only forms in *u* vocalism are reconstructed for Proto-Eastern-Lezgian, there is no

Lezgian subgroup, the distinction between absolutive and ergative has been lost in the independent personal pronoun in all dialects of Tabassaran and in most dialects of Aghul. As described above, the distinction has also been drastically reanalyzed in the system of verbal suffixes of the northern dialect of Tabassaran¹³

4.3. Completing the Picture. On the basis of the facts discussed above, I postulate the following sort of development. In Proto-Eastern-Lezgian pronouns, like nouns, had distinct forms for absolutive and ergative. After the break up of this subgroup, Tabassaran innovated the practice of optionally copying topicalized pronouns as enclitics to finite verb forms. It was specific case forms that were copied, so that ergative case subjects were copied as forms with vocalism *a*, while absolutive case subjects were copied as forms with vocalism *u* (see examples (9) and (10)). Subjects were most frequently topics and copy-cliticization became obligatory, thus constituting agreement. In some dialects it remained possible to copy-cliticize topics that were not subjects in the position following the agreement suffixes (see example (11)).

Since obligatory subject agreement distinguished two forms, those in *u* and those in *a*, it combined aspects of ergative alignment with aspects of accusative alignment. The accusative alignment of agreement was reflected in the fact that subjects of both transitives and intransitives alike were triggers of obligatory agreement. The ergative alignment of agreement was reflected in the fact that direct objects of transitives, which were optionally copied onto the verb form, and subjects of

¹³Dixon (1979:92) has suggested that, since agreement ("cross-referencing systems") usually develop from pronouns, and since pronouns are basically (nominative-)accusative in alignment on the animacy hierarchy, we expect agreement also to be accusative in alignment, at least at first. There is little doubt that first and second person personal pronouns in Proto-Eastern-Lezgian (and in Common Lezgian) were, however, not accusative in alignment, but ergative(-absolutive). Thus, while Dixon's observation does account for the development of accusative agreement in a hypothetical language that has ergative case marking for nouns and accusative case marking for pronouns, it does not explain the Tabassaran development. For the same reason, Dixon's observation alone does not account for the apparent universal that accusative agreement may occur in a language with ergative case marking, while the reverse does not occur.

intransitives were marked in a form that was distinct from that used to mark subjects of transitives. The ergative and absolutive of the first and second person independent pronouns fell together, continuing only the form that earlier was absolutive; ergative and absolutive forms of agreement markers remained distinct. This stage is preserved in the southern dialect.¹⁴ The northern dialect made the following additional innovation: the ergative and absolutive of the first and second person agreement markers fell together, continuing only the form that earlier was ergative. The forms in *u*, which had originally represented all absolutive nominals, were restricted to the optional clitic-copy of direct objects in the presence of a non-null marker of subject agreement. Only at this stage, did person agreement become a fully accusative rule, triggered by subjects of both transitive and intransitive, which condition markers of the same form.¹⁵

5. Conclusion. It has been proposed here that the development of agreement as a rule with accusative alignment in Tabassaran was crucially dependent upon the following factors:

1. loss or weakening of the inherited system of grammatical class
2. copying of topic pronouns as verbal clitics
3. high incidence of topics as subjects
4. syncretism (secondarily developed) of ergative and absolutive case in the forms used as clitic-copies.

This does not deny the possible importance of the influence of contact languages in encouraging the development of accusative agreement,

¹⁴ Batsbi, a distantly related language of the Nax subgroup of North East Caucasian, has an agreement pattern similar to that reconstructed here for pre-Tabassaran, in that it preserves inherited class agreement prefixes assigned on an ergative pattern, combining them with person agreement suffixes that vary according to the case of the subject, which they encode. For examples, see Holisky, to appear, sections 2.5.1.1 and 2.5.1.2. I am grateful to Johanna Nichols for drawing my attention to this similarity.

¹⁵ The southern dialect differs slightly from the reconstructed situation in that (according to Kibrik 1979:74-75) in that dialect subjects of active intransitives are in the form reconstructed here as ergative. It is not possible to determine at present whether this feature is a dialect-specific innovation or whether it represents the original situation. Given this indeterminacy, I have referred to it as ergative for the sake of simplicity.

but seeks to account for why and how the innovation was made from a language-internal point of view.

This hypothesis offers an explanation for the otherwise unexplained fact that in the Lezgian languages, innovative agreement with accusative alignment occurs only as agreement in person and number, while the inherited ergative system indicates agreement in gender-class. This hypothesis traces the stages of the ergative-to-accusative shift in agreement through evidence remaining in various dialects.¹⁶

REFERENCES

- Alekseev, M.E. 1985. *Voprosy sravnitel'no-istoričeskoj grammatiki lezginiskix jazykov*. Moscow: Nauka.
- Anderson, Stephen R. 1977. On mechanisms by which languages become ergative. *Mechanisms of syntactic change*, ed. by Charles Li, 317-363. Austin: University of Texas.
- Andžulage, N. 1968. K'lasovani da p'irovani uŷvlilebis ist'oriis zogi sak'itxi iberiul-k'avkasiur enebši. [Some questions of the history of class- and person-conjugation in the Ibero-Caucasian languages.] Tbilisi: Mecniereba.
- Čikobava, Arn. 1937. Gramat'ik'uli k'lasebis ist'oriisatvis xunžurši. *Enimk'is Moambe* 1:97-105.
- Čikobava [Cikobava], Arnold. 1962. Les types fondamentaux de la conjugaison des verbes et leurs relations historiques dans les langues ibero-caucasiennes (ou caucasique). *Bedi Kartlisa* 13-14:26-33.
- Chung, Sandra. 1978. Case marking and grammatical relations in Polynesian. Austin: University of Texas.
- Dešeriev, Ju. D. 1959. *Gramatika xinalugskogo jazyka*. Moscow: Akademija.
- Dixon, R.M.W. 1979. Ergativity. *Language* 55:59-138.
- Du Bois, John W. 1985. Competing motivations. *Iconicity in syntax*, ed. by John Haiman, 343-365. Amsterdam: Benjamins.
- Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63:805-855.
- Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. *Subject and topic*, ed. by Charles N. Li, 149-188. New York: Academic Press.
- Harris, Alice C. 1984. Case marking, verb agreement, and inversion in Udi. *Studies in Relational Grammar*, 2, ed. by David M. Perlmutter and Carol G. Rosen, 243-258. Chicago: The University of Chicago Press.
- _____. 1985. *Diachronic syntax: The Kartvelian case*. New York: Academic Press.
- Holisky, Dee Ann. To appear. Tsova-Tush (Batsbi). *The indigenous languages of the Caucasus*, 3: North East Caucasian, ed. by Michael Job and Rieks Smeets.

¹⁶The discussion of rule alignment does not take into account agreement in the inversion construction. When these verb forms are considered, it can be seen that even in the northern dialects agreement is still not fully accusative.

- Delmar, N.Y.: Caravan Press.
- Ĵavaxišvili, Iv. 1937. Kartuli da k'avk'asiuri enebis tavadp'irveli buneba da natesaoba. [The original nature and genetic affiliation of Georgian and the Caucasian languages.] Tbilisi: Ak'ademia.
- Ĵeiranišvili, Evgeni. 1953. Gramat'ik'uli k'lasebi c'axurula da rutulur enebši. Ibero-k'avk'asiuri enatmecniereba 5:451-469.
- _____. 1956. Gramat'ik'uli k'lasis gakhvavebuli nišnebi udur zmnebsa da nazmnar saxelebši. [Fossilized markers of grammatical class in Udi verbs and deverbal nouns.] Ibero-k'avk'asiuri enatmecniereba 8:341-362.
- _____. 1971. Udiuri ena [The Udi language]. Tbilisi: Universit'et'i.
- Kaxaze, Ot'ia. 1979. Arčibuli ena da misi adgili dayestnis monatesave enata šoris. [The Archi language and its place among the related languages of Daghestan.] Tbilisi: Mecniereba.
- _____. 1983. Gramat'ik'uli k'lasebis taobaze lezgiur enebši. C'elic'deuli 10:319-333.
- Kibrik, A.E. 1979. Canonical ergativity and Daghestan languages. *Ergativity*, ed. by Frans Plank, 61-77. New York: Academic Press.
- _____. and M.G. Seleznev. 1982. Sintaksis i morfologija glagol'nogo soglasovanija vtabasaranskom jazyke. *Tabasaranskije ètjudy*, ed. by G.A. Klimov and B.Ju. Gorodeckij, 17-32. Moscow: Universitet.
- Klokeid, Terry J. 1978. Nominal inflection in Pamanyungan: A case study in relational grammar. Valence, semantic case, and grammatical relations, ed. by Werner Abraham, 577-615. Amsterdam: Benjamins.
- Magometov, A.A. 1965. *Tabasaranskij jazyk*. Tbilisi: Mecniereba.
- _____. 1970. *Agul'skij jazyk*. Tbilisi: Mecniereba.
- Mallinson, Graham, and Barry J. Blake. 1981. *Language typology*. Amsterdam: North Holland.
- Pančvize, Vl. 1974. Uduri enis gramat'ik'uli analizi [A grammatical analysis of the Udi language]. Tbilisi: Mecniereba.
- Plank, Frans. To appear. Syntactic change: Ergativity. *Syntax: An international handbook of contemporary research*, ed. by Joachim Jacobs, Arnim von Stechow, and Theo Vennemann. Berlin: de Gruyter.
- Rogava, G. 1962. Zmnis p'olip'ersonalizmis ist'oriisatvis iberiul-k'avk'asiur enebši. [On the history of polypersonalism of the verb in Ibero-Caucasian languages.] *Iberiul-k'avk'asiuri enatmecniereba* 13:45-50.
- Sapir, Edward. 1917. Review of *Het passieve karakter van het verbum transitivum of van het verbum actionis in talen van Noord-Amerika*, by C.C. Uhlenbeck, 1916. In *International Journal of American Linguistics* 1:82-86.
- Schulze, Wolfgang. 1982. *Die Sprache der Uden in Nord-Azerbajdan*. Wiesbaden: Harrassowitz.
- Topuria, G. 1960. Zogierti c'inasc'ari cnoba k'ric'ul enis morfologiidan. [Some preliminary information on the morphology of the Kryts language.] *Iberiul-k'avk'asiuri enatmecniereba* 12:437-442.

- _____. 1962. K'lasovani uɣvlilebidan p'irovan uɣvlilebaze gadasvlis ert-erti gzis šesaxeb iberiul-k'avk'asiur enebši. [On one of the routes of transition from class conjugation to person conjugation in the Ibero-Caucasian languages.] *Iberiul-k'avk'asiuri enatmecniereba* 13:55-65.
- Xajdakov, S.M. 1980. *Principy imennoj klassifikacii v dagestanskix jazykax*. Moscow: Nauka.
- Žirkov, L.I. 1948. *Tabasaranskij jazyk*. Moscow-Leningrad. [Not seen.]

ON A STRUCTURAL 'FIFTH COLUMN' IN SOCIOLINGUISTIC CHANGE: THE DIFFUSION OF A STANDARD YIDDISH FEATURE IN YIDDISH DIALECTS

Neil G. Jacobs
The Ohio State University

1. Introduction.¹ The development of a modern Standard Yiddish (StY) language in Eastern Europe does not show a uniform linear history. Until approximately 1800, a written standard based on Western Yiddish (WY) written norms generally prevailed throughout the Yiddish speech territory—both west and east—even as the modern Eastern Yiddish (EY) spoken dialects had already largely crystallized (not later than approximately 1700 (M. Weinreich 1980:733)). Beginning early in the 19th century, however, a standard language based on EY speech norms began emerging. Standardization trends continued into the 20th century, and the acceptance of StY (in somewhat varying forms) was greatly furthered and accelerated by the emergence in the 1920s of three Yiddish institutes in Eastern Europe with language-academy function: in Vilne (Lithuanian *Vilnius*), in Minsk, and in Kiev.

We must not limit ourselves to a view of standardization trends as a unidirectional endpoint, that is, from dialect to standard language. Many of the changes in EY dialects in the 20th century may best be viewed in terms of an interaction between dialect and standard, with one or more features from the standard becoming incorporated and integrated into dialect. One well-known example of this is the restoration of the hushing/hissing distinction for sibilants in Northeastern Yiddish (NEY), which occurred under the influence of emerging StY; see U.

¹ A preliminary version of this paper was presented at the Seventh Conference on the Non-Slavic Languages of the Soviet Union, May 22-24, 1991, at the University of Chicago. Discussion of glides in Yiddish is based on my paper (Jacobs 1990/91) presented at the Linguistic Society of America Annual Meeting, held in Chicago, January 3-6, 1991. I am grateful to Robert P. Stockwell (UCLA) for a number of insightful comments and suggestions which contributed to the discussion in the current paper. I also wish to thank the following units at The Ohio State University for their generous research and conference support: the Department of Judaic and Near Eastern Languages and Literatures, the Melton Center for Jewish Studies, the Center for Slavic and East European Studies, and the Middle East Studies Center. Naturally, any and all shortcomings and/or mistakes in the present paper are the responsibility of the author.

Weinreich 1952.²

The present paper seeks to examine a specific example of the interaction between dialect and standard. Specifically, this paper offers a discussion of the linguistic nature and geographic extent of the diffusion of a StY diphthong in EY dialects. The data and discussion are based on the earlier findings of U. Weinreich (1958). Weinreich was here primarily concerned with the realization of Proto-Germanic (P-Gmc) **u*: in EY dialects. In two geographically non-contiguous subregions within NEY and Southeastern Yiddish (SEY), respectively, this vowel is realized as a monophthong *u*(:). In Weinreich's important paper it was shown that these monophthongs were the results of innovation, rather than historical retentions. Weinreich's arguments were based on consideration of the internal phonemics of the vowel systems of each of the two subregions. Thus, Weinreich was chiefly concerned with *why* the two *u*-regions existed, rather than with the more recent diffusion of a StY feature. This is reflected in Weinreich's (1958:225) statement: "Unfortunately for the progress of the argument, the facts about the occurrence of *u* < **ū* in present-day Yiddish are complicated by dialect leveling and the spread of Standard Yiddish."

The present paper is concerned with the *where* and *how* of the diffusion of the StY *oj* cognate with P-Gmc **u*: (the vowel dealt with by Weinreich 1958; to be referred to as vowel 54, see below). Generally, it may be said that this StY *oj* diffused into NEY and SEY areas, but not into Central Yiddish (CY) for reasons that are intimately related to Weinreich's findings.³ It will be shown that this StY diphthong diffused in different ways in the areas into which it did spread. Specifically, StY *oj* (vowel 54) diffused phonologically in NEY, but lexically in SEY. The differing nature of diffusion in the two areas will be shown to result from the differing diachronic phonologies of the two dialects.⁴

²On the influences of the 19th century phenomenon of *daytshmerish* (the heavy influence from New High German) upon StY, see Schaechter 1969. For recent discussion of Yiddish standardization processes see Hoge (1991).

³WY will not be considered in this discussion, since it was virtually out of the picture as concerns the developments under discussion in the present paper.

⁴Although these structural considerations are shown to be relevant in the present paper, no claims are made here regarding the debate between Jakobson's (1938) claim that languages/dialects will only accept an external influence if they are structurally "ready" to accept such influence, versus the strong caution against claims of universality found in Thomason and Kaufman (1988).

2. The vowel in question: PY vowel 54.

2.1. Reconstruction and the PY vowel numbering system. The vowel under present discussion is reconstructed as P-Gmc **u:*. This P-Gmc vowel has undergone diphthongization in most Gmc daughter dialects except for those at the geographic periphery; thus, P-Gmc **hu:s-* 'house,' **mu:s-* 'mouse' > English *house, mouse* ([aw]), German *Haus, Maus*, Dutch *huis, muis*, whereas North Germanic and Swiss German dialects generally have preserved some sort of long high monophthong (e.g., Swedish *hus* ([*ʉ:*])).

The present paper will follow M. Weinreich's (1960) commonly-employed numbering system as modified in Herzog (1965), and refer to the vowel under discussion as vowel 54. In this system, the first digit refers to the original Proto-Yiddish (PY) quality of the vowel: [a] = 1-; [e] = 2-; [i] = 3-; [o] = 4-; [u] = 5-. The second digit refers to PY length or diphthongization: -1 = short; -2 = long; -3 = originally short, lengthened early on,⁵ -4 = diphthong, and -5 (= special notation, used solely to distinguish PY **e:22* — **ei24* — **æ:25*. Max Weinreich (1980:674) reconstructs PY diphthongs 34 and 54 as **iit* and **uut* respectively. However, the plausibility of these being phonetically distinct from PY long monophthongs **i:32* and **u:52* may be questioned. Furthermore, there is evidence to suggest that the nuclei of PY diphthongs 34 and 54 were already [-high] (that is, either [əj]₃₄, [əw]₅₄ or [aj]₃₄, [aw]₅₄) by PY times, since substratal Hebrew-Aramaic (HA) component *i:* and *u:* did not take part in the diphthongization.⁶ Thus: HA-origin **gəvu:ɾɔ:* 'courage,' **gəvi:ri:m* 'rich men' > PY **gvu:ɾɔ*, **gvi:rim*, with monophthongs *i:* and *u:*, and not with diphthongs (see comment on Wexler 1991, in Jacobs 1991:178f.). Thus, with our modification (PY **aj34*, **aw54*), the PY vowel system may be given as follows:⁷

⁵The historical validity of the -3 series has come into question in more recent works; see Katz (1982); Jacobs (1990).

⁶For a differing reconstruction, see U. Weinreich 1958:232, 247.

⁷We may ignore here Max Weinreich's -3 series ("early lengthened"), since—even if historically valid—this distinction is not maintained in EY dialects. The phonetic values posited for [æ:]₂₅ and [â:]₁₂ are taken from Jacobs 1990.

| Short | | Long | | Diphthongs | |
|-----------------|-----------------|------------------|------------------|------------------|------------------|
| i ₃₁ | u ₅₁ | i: ₃₂ | u: ₅₂ | | |
| e ₂₁ | o ₄₁ | e: ₂₂ | o: ₄₂ | ej ₂₄ | ow ₄₄ |
| a ₁₁ | | æ: ₂₅ | ǣ: ₁₂ | aj ₃₄ | aw ₅₄ |

In the discussion below, M. Weinreich's numbering system will be employed, since it allows us to make unambiguous reference to the diaphonemes in the modern Yiddish dialects, as well as to StY. In this way, for example, we may refer to StY *oj*₅₄ (< PY **aw*₅₄), as distinct from StY *oj*₄₂ (< **o*:₄₂) and StY *oj*₄₄ (PY **ou*₄₄).⁸

2.2. Some sample words with StY *oj*₅₄ are:

| | | | |
|--------------|-----------|---------|---------------|
| hojz | 'house' | mojz | 'mouse' |
| hojt | 'skin' | (ar)ojz | 'out' |
| (vajn)trojbn | 'grapes' | tojznt | 'thousand' |
| krojz | 'cabbage' | jojz | 'broth, soup' |

2.3. EY realizations of vowel 54. The common realizations of vowel 54 in the major modern EY dialects are as follows (see U. Weinreich, 1958:230, map 2):

| | | |
|---------|------------------|-----|
| CY | NEY | SEY |
| ou ~ o: | oj ~ uj ~ u: ~ u | u |

The two *u*(:)-areas—one in NEY, the other in SEY—are separated by large areas with [oj] (*oi*) and [uj] (*ui*).¹⁰

The 20th century has seen an ongoing diffusion of StY *oj*₅₄ outward

⁸U. Weinreich's (1958:226) designation of PY vowel 54 as "vowel 15" is idiosyncratic to that article, and will not be used in the present paper.

⁹Yiddish *jojz* ultimately derives from Slavic *jûxa*. However, a German intermediary source with diphthong is assumed; cf. German *Jauche* 'sewer waste.'

¹⁰See Weinreich (1958:230, map 2). The recently published *LCAAJ*, map 37 (page 86) shows a southern *u*-area, pockets of *uj*, but not the NEY *u*-area found in Weinreich (1958). Thus, this would appear to present a challenge to Weinreich's earlier (northern) *u*-area. However, even in Weinreich (1958), the northern *u*-area was geographically very limited, and had long been undergoing a process of giving way to the diffusing *oj* under discussion in the current paper. Thus, for purposes of our current discussion, it is sufficient to note the historical presence of the northern *u*-area.

into EY dialects. For the most part, however, StY *oj*₅₄ has diffused into the NEY and SEY areas, but not into the CY area. We are impelled to ask why this is so.

3. Vowel length, the glide *w*, and the diffusion of StY *oj*₅₄. The claim is made here that the diffusion of StY *oj*₅₄ was generally limited to areas lacking synchronic distinctive vowel length (i.e., NEY and SEY). StY *oj*₅₄ generally did not diffuse into areas with distinctive vowel length (CY, as well as the conservative Courland subregion of NEY).¹¹ As will be shown below, however, the issue of distinctive vowel length by itself does not "explain" the limits on diffusion of StY *oj*₅₄. Rather, it is another issue—the relationship of vowel length and glide types in Yiddish dialects—which will help clarify the matter at hand.

In Jacobs (1990-91) a direct relationship was shown between the presence of distinctive vowel length and the presence of a rounded glide in Yiddish dialects. WY and CY—both possessing distinctive vowel length—have a back/round glide *w* (in addition to front/unround glide *j*). Most varieties of NEY and SEY generally lack distinctive vowel length. There is arguably marginal residue of length for a single vowel quality in each area: *u*: in a subregion of NEY, and *i*: in parts of the SEY area. However, these may be seen as the final—and controversial—vestiges of length in now lengthless systems. Most varieties of NEY and SEY generally lack the back/round glide *w* as well.¹² Interestingly, U. Weinreich (1958:260) relates the monophthongization of vowel 54 to loss of vowel length in SEY: "In the Southeast, the process was due to the elimination of distinctive vowel quantity, presumably under Slavic influence." However, the connection between vowel length and glide type is not stated here by Weinreich.

The synchronic situation in the major Yiddish dialects may be represented as follows:

¹¹On vowel length in Courland Yiddish, see M. Weinreich 1923, Kalmanovitsh 1926, Jacobs (forthcoming). U. Weinreich (1958:262) notes that StY *oj*₅₄ did diffuse into territory that was "previously" CY, i.e., precisely into areas that had—as a result of dialect contact—developed a vowel system lacking the length feature.

¹²There is some evidence to suggest that of the two features—[+back] and [+round]—it is [+round] that is crucial. This claim is based on the presence of diphthong [øy]—with a front/round glide—in Courland Yiddish (CourlY) subregion of NEY.

The realization of vowel 54 as [oj] is seen as originating in the area around Vilne, and to have spread from there (U. Weinreich 1958:227). Vilne-type speech was seen as a prestige variety, and played a significant part in the pan-EY standardization trends of the 19th and 20th centuries. Thus, the motivation for the borrowing of *oj*₅₄ in Yiddish dialects was primarily sociolinguistic—related to the prestige associated with a Vilne-type pronunciation. In the present paper, “Vilne type” or “Vilne variant” is employed to mean the prestige speech of Vilne and its surroundings. It was in this speech that *oj*₅₄ originated, and whence it diffused into StY as well as into surrounding dialects. U. Weinreich (1958:249) makes additional reference to two other versions of the “Vilna system”—the Mogilev variant, with raised nucleus and front glide: [o̞̠ , u̞̠], and the Vitebsk variant, with back glide: [o̞̠̠ , Uu̞̠̠ , u̞̠̠]. In the present paper, “Vilne type” will be used to refer exclusively to the *oj* variant which went from Vilne speech into StY.

As regards the diffusion of StY *oj*₅₄ into the NEY and SEY areas, and its non-spread into the CY area, we should consider briefly the following simple argument. It could be claimed that StY *oj*₅₄ diffused only into areas where it would not undermine a phonemic distinction with a preexisting /oj/. In this view, StY did not diffuse into CY because of a preexisting CY *oj*_{42,44}; thus, the phonemic distinction in CY *tojb*₄₄ 'deaf' ([tojp]) vs. CY *to:b*₅₄ ([to:p]) 'dove' would be lost

¹⁴Vilne-area variety, which served as a prestige model in diffusion of *oj*₅₄; see below.

through introduction of StY *toj*₅₄. This argument fails, however, in light of the diffusion of StY *oj*₅₄ into the SEY area, which had preexisting *oj*_{42,44} as well.¹⁵ Furthermore, synchronic StY shows as well just this full merger of vowels 42, 44, and 54 as StY *oj*_{42,44,54}. Thus, the limits on diffusion of StY *oj*₅₄ are not based solely on the issue of the phonemic distinctiveness of /oj/ in the EY dialects. Rather, these limits are linked to deeper phonological issues of vowel length and the status of glides in Yiddish dialects.

4. Data and analysis.

4.1. Vowel 54 in EY: Evolution and chronology. The reconstructed PY and Proto-Gmc values for Yiddish vowel 54 were discussed above. Our current discussion is limited to EY. The relevant issue is that we must assume that vowel 54 was a diphthong in Proto-EY (PEY). I suggest PEY **aw*, based on data in U. Weinreich's article. Weinreich's two separate EY *u*(:) regions may be seen as developing through historical processes of: (1) raising, (2) monophthongization, and (3) (regional) shortening:

**aw* > **ow* > *uw* > *u*: (> *u*)

In both NEY and SEY there occurred a hardening of *w* → [v] in intervocalic position. When hardening occurred, the raising process in the nucleus was halted. Thus, while vowel 54 could continue raising to *u*(:) / _ C (e.g., [hut]₅₄ 'skin'), the raising process was halted after intervocalic hardening (e.g., NEY regional *zavər* and SEY *zovər*₅₄ 'sour'; see now *LCAAJ* map 40, page 89). P(E)Y **aw*₅₄ is posited in order to account for NEY regional variant *zavər*₅₄ 'sour.' Such regional variants suggest the chronologically early appearance of intervocalic hardening, prior to the onset of any raising process. Glide hardening is found (though not uniformly) in the modern reflexes of other historical vowels as well in NEY and SEY: widespread NEY *evər* 'ear' (see *LCAAJ* map 31, page 80) < PY **o*:₄₂, cf. StY *ojər*. For NEY *evər* we assume historically intermediate **øyər* or **øwər*.

On the face of things, both NEY and SEY appear to exhibit considerable similarity, both in their general evolution of vowel 54 (from **au* > *u*(:)), and in the process of intervocalic hardening of glide

¹⁵The status of this diphthong in SEY is somewhat problematic. See Vilenkin 1931: Maps 42-49.

$w > [v]$. The phonological roles of the two processes were different, however, in NEY and SEY.

4.2. Diphthongs, vowel length, and glide type. As mentioned above, all Yiddish dialects possess surface diphthongs. However, in Jacobs (1990-91) it was pointed out that those Yiddish dialects with distinctive vowel length possess both glides, j and w , whereas those Yiddish dialects lacking distinctive vowel length have glide j , but no glide w . It may be asked what the relationship is of the w glide—but apparently not the j glide—to vowel length? Consider the following examples, given in StY and CY forms:¹⁶

| <u>StY</u> | <u>CY</u> | <u>gloss</u> |
|------------|-------------|-----------------------------|
| ojg | [oj] | 'eye' |
| jojx | [ow] ~ [o:] | 'soup, broth' ¹⁷ |

However, both length- and lengthless-varieties of Yiddish lack initial $[w]$; thus, there is no: ***wojx*, ***wojn*, etc.

In these examples the glide j may occur word- (syllable- ?) initially, whereas the glide w may not. Stated otherwise: at least in initial position the glide j may function as a consonant, whereas the glide w may not.¹⁸ If this analysis is extended to our current discussion of diphthongs, then it is possible to interpret those diphthongs with j glide (*aj*, *oj*, *ej*) as:

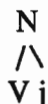


¹⁶It was correctly pointed out to me by an anonymous reader that merely giving StY examples here would not be sufficient, since (lengthless) StY lacks the $[w]$ glide. The reader wrote: "...it would be better to give examples from dialects with the glide $/w/$ also, so that one could more clearly see the lack of symmetry." I am grateful to the reader for this observation.

¹⁷The variants and geographic distribution of this lexical item are problematic for other reasons; see LCAAJ map 38, page 87. For the typical development of this vowel (vowel 54), see LCAAJ map 37, page 86, with words such as *mojz* 'mouse' (CY [ow] ~ [o:]).

¹⁸From a phonemic perspective, U. Weinreich (1958:260) writes: "In NE Yid. remonophthongization resulted rather from the instability of the only diphthong with a u -offglide in a dialect without a u phoneme."

that is, a sequence of vowel + consonant, where consonant *j* is part of the coda, rather than the nucleus. Alternatively, the glide *j* could be analyzed as [+vocalic], and part of a branching (= [+long]) syllable nucleus:



Thus, two possible interpretations are open for the glide *j*. On the other hand, if we say that the glide *w* may function only as a vowel, then surface diphthongs with *w* could only be interpreted as containing a branching/long syllable nucleus, since [+vocalic] *w* must be part of a nucleus, and not part of a coda:



Long monophthongs may also be represented as branching syllabic nuclei:



In this analysis, diphthongs with glide *w* are viewed as similar to long monophthongs in that both contain a branching/long nucleus. On the other hand, it was possible to analyze diphthongs with glide *j* as sequences of short vowel + consonant; that is, short/non-branching nucleus followed by non-nuclear consonant *j*. In this analysis, the presence of glide *w* requires recognition of vocalic length, whereas the glide *j* does not. From this perspective, it makes sense that both major Yiddish dialect areas lacking systematic long vowels—NEY and SEY—have developed various strategies to eliminate [w], but not [j]. The fact that NEY (earlier) and SEY (later) lost vowel length at different times has led to divergent development in the relevant data. We thus are led now to examine the NEY and SEY developments separately.

4.3. NEY. Loss of long vowels is posited as a primary event in the evolution of NEY (see, e.g., Herzog 1965:197ff).¹⁹ The early loss of length led to the systematic merger of all long/short monophthongal distinctions inherited from PEY, as follows:

| PEY | | NEY |
|-------------------------------------|---|--------------------|
| *i ₃₁ ≠ *i ₃₂ | > | i _{31,32} |
| *e ₂₁ ≠ *e ₂₅ | > | e _{21,25} |
| *o ₄₁ ≠ *o ₁₂ | > | o _{12,41} |
| *u ₅₁ ≠ *u ₅₂ | > | u _{51,52} |

The relationship of the loss of the *w* glide to the loss of vowel length in Yiddish dialects was claimed above. Accordingly, NEY shows general elimination of the *w* glide (elimination thus not limited to vowel 54). Several independent processes have conspired to produce this common elimination of glide *w*. They are:

(A) Intervocalic hardening: **zawər* / **zowər* > *za[v]ər* / *zo[v]ər* ‘sour’; NEY *mevax* ‘brain,’ *kevax* ‘force’ < PEY **mowax*, **kowax* < HA-origin (Tiberian Hebrew) *mo:ax*, *ko:ax*.

(B) Positional deletion: *w* is deleted before labial(ized) consonants:

| | | | |
|---------------|---|------------|----------|
| * <i>kawm</i> | > | <i>kam</i> | ‘hardly’ |
| * <i>mowt</i> | > | <i>mot</i> | ‘mouth’ |

(C) (Sporadic) hypercorrection, based on heavily labialized NEY *ʔ*. This occurs /_ C; e.g., **krowt* > *kroʔt* ‘cabbage’; cf. StY *krojʔ*.²⁰

(D) Deletion of last resort: If (A), (B), and (C) above are not available, then simply drop the *w*: **blow* > *blo* ‘blue’; cf. StY *bloj*.²¹

¹⁹The notable exception here is the Courly subregion, where vowel length has remained into the present century.

²⁰Weinreich (1958:230), map 2, obtains his *ul* area from Veynger 1926, and Veynger 1927-28. See also *LCAAJ*, map 58, page 107.

²¹Note, however, the inflected forms *blowə*, *blowər*, etc. < **blowə*, **blowər*, with original intervocalic *w*. Furthermore, morphological paradigm leveling has also led to the presence of *v* in the base form of adjectives, e.g., NEY *rev* ‘raw’

(E) A fifth strategy for the elimination of the *w* glide in NEY is that of Glide Substitution (GS) (Jacobs 1990-91), where the glide *j* replaces round glide *w* (and *y*). The process of GS is limited to the environment /_ C. Thus, PEY **towrə* (vowel 42) > Proto-NEY **tøyrə* > NEY *tejrə* 'Torah.' (Intervocalically, *w* hardens to [v]: Proto-NEY **øwər* (vowel 42) > NEY *evər* 'ear'.²²

Historically, GS is assumed to have occurred during a period when intervocalic hardening had not run its course, since GS generally affects sequences of *w* + C, but not sequences of /v/ + C²³; thus:

**towrə* > *tøyrə* > NEY *tejrə* 'Torah'
 xevrə* > NEY *xevrə*, not *xejrə* 'gang'

There are examples—limited, few in number, and regionally sporadic—in which the sequence *v* + C has undergone hypercorrection / reanalysis based on NEY GS; thus: *baha[v]nt* 'well-versed,' *davnən* 'to pray' have yielded hypercorrect *baho[j]nt*, *do[j]nən*; see Weinreich (1958:243, footnote 16). Note, however, that the hypercorrect application of GS could only occur where the structural description of GS was met, since the [v] did not undergo GS > *j* intervocalically. Thus, the alternate form of 'to pray'—*davənən*—was not eligible to undergo hypercorrect GS to ***dojənən*.

All the processes listed above (A-E) have formed a conspiracy, sharing the single net effect of eliminating the *w* glide. Each process thus helped eliminate potential long/branching nuclei from a dialect which was ridding itself of long vowels. GS fits into this line of argumentation, however, only if we go along with the analysis of glide *j* as consonantal in those Yiddish dialects which lack systematic vowel length.

(PNEY **røɣ* < PEY **row*₄₂), based on NEY inflected forms *revə*, *revər*, etc.; see U. Weinreich 1958:253.

²²On regional NEY forms *ejər* and *ojər*, see below.

²³Apparent counterexamples are the Jewish names recorded in Russian documents (U. Weinreich 1958:253): *Jevno* 'Jonah,' *Pevsner*, etc. These show fricative <*v*> rather than a glide—either *w* or *j*. However, these examples may reflect adaptation to lexicalized Russian phonological structure (e.g., /v/ in Russian *avgust* 'August,' *avtomobil* 'automobile,' versus glide in other European languages, including Yiddish).

The view of GS as phonological in nature is supported by consideration of another NEY process which may be called Glide Insertion (GI). In GI, *j* is inserted as a hiatus breaker between adjacent vowels, thus:

$$V_i V_{ii} \rightarrow V_i j V_{ii}$$

Consider the development in the following HA-origin words from spoken-Hebrew times to modern NEY:

| HA | PY | NEY | gloss |
|------------|---------------------|----------|-----------|
| *tɔ:fʊ:θ > | *tɔ:əs > /toəs/ → | [tojəs] | 'mistake' |
| *rəfu:ʔ > | *rəfu:ə > /rəfuə/ → | [rəfujə] | 'remedy' |

Insertion of consonant *j* thus breaks up sequences of vowels. Note that the occurrence of GI in NEY [rəfujə] provides evidence for the historically early loss of long vowels in NEY. If the original HA *u:* had retained its length historically too late, then intervocalic hardening would have yielded ***rəfuvə*.

Taken together, GS and GI suggest that NEY speakers used the *j* glide as an available phonological strategy to eliminate vocalic sequences.

Consider again the main variants of vowel 54 in NEY:

[oj] [uj] [u:] [u]

I offer the following account for the NEY regional variants, taken as an interconnected dialectological set. The *oj* variant developed in the Vilne area. In this area the historical raising of the nucleus of vowel 54 had halted at mid-vowel level, thus yielding **ow*₅₄. Perhaps not coincidentally, this is identical at this stage to (geographically adjacent) CY *ow*₅₄. I then suggest that the loss of vowel length originated (for NEY) in this Vilne region, triggering the elimination of the *w* glide for reasons discussed in the current paper. Vowel length remained on the historical scene longer, however, in the NEY areas further east (Mogilev and Vitebsk areas). In these eastern areas, the continued presence of the *w* glide permitted the further raising of the nucleus of vowel 54 from *ow* > *uw*. The loss of phonemic vowel length then spread to the eastern areas of NEY (under Vilne-type or standardizing influence?).

At this point, a fully-raised uw_{54} —or $u:_{54}$ —presents a problem of anomalous length in a now lengthless system. Two main approaches to deal with this problem of anomalous length present themselves:

- (a) Analyze as a diphthong, and apply GS: $uw > uj$.
- (b) Analyze as a long monophthong, and simply shorten it: $u: > u$.

A third possibility, and one that does occur, is that the evolution of the lengthless system had not run its full course, and long $u:$ remains as a residue of earlier length. This would seem to be the case in one subregion of NEY. Interestingly, however, the uj subregion intervenes geographically between the area with Vilne/StY oj (to the west) and the long $u:$ area (to the east).²⁴ This suggests that GS was indeed a strategy employed to deal with the problem of anomalous vowel length in a lengthless system. Furthermore, the uj area is geographically homogenous, and applies generally—that is, phonologically rather than lexically—in this subregion. It might be asked (U. Weinreich 1958:256, N. 28, citing Birnbaum 1934:28n.) if uj represents a type of dialect-contact hybrid form, i.e., $oj + uw \Rightarrow uj$?) This possibility might not be ruled out. Consider the NEY dialect form $evər$ 'ear' < $*\emptyset yər$ < $*owər$. It is possible that the NEY variant $ejər$ 'ear' arose as a contact-induced compromise between dialect form $evər$ and StY $ojər$. This is supported by the fact that one would not regularly expect the application of GS intervocalically.

Finally, the instability of long $u:/uw$ in the now lengthless subregion has also led in some locations to the development $u:/uw > [ul]$. Weinreich (1958:256) attributes the development of $[ul]$ as follows: (1.) realization of $/l/$ positionally as $[w]$ (via contact with coteritorial Belorussian);

²⁴In Weinreich (1958:230, map 2), the $[uj]$ (ui) subregion is shown graphically as a homogeneous area, bounded by u -areas to the north and south, and an $[oj]$ (oi) area to the west. However, no toponyms are provided in this map. Weinreich's (1958:261) map 5—which gives a diachronic view of the geographic variants of vowel 54, and the direction(s) of diffusion of $[oj]$ (oi)—does provide toponyms. The once-larger ui -area, as shown in map 5, was bounded, roughly, by Kaunas in the northwest, proceeding eastward past Mogilev in the northeast, above Zhitomir and Kiev in the south, and including Brest Litovsk and Grodno in the west. Further, map 5 shows the sub-area—roughly bounded by Vilne, Pinsk, Brest Litovsk, and Grodno—as the source of innovation and subsequent outward diffusion of $[oj]$ (oi).

(2.) subsequent replacement of all [w]—i.e., etymological *w* as well as etymological *l*—with lateral [l] under influence of dialect leveling and/or standardization.

4.4. SEY

4.41. Lexical vs. phonological diffusion. Discussion of SEY will be brief, since most of the issues have been laid out in terms of NEY, above. Historically, SEY lost vowel length relatively late.²⁵

The diffusion of StY *oj* into SEY territory has been lexical rather than phonological. U. Weinreich (1958:229) describes this diffusion as "...in the southern area...word by word replacement..." This is in contrast to the situation of phonological generalization in NEY: "In the north, the isoglosses are much less staggered" (U. Weinreich 1958:229). On the lexical diffusion of StY *oj* into SEY, U. Weinreich (1958:233) provides the following examples from one SEY speaker (from Bar, born approximately 1885):

| | | |
|-------------------------------|-------------------|--------------|
| Diffusion of StY diphthong: | <i>oiz</i> | 'house' |
| | <i>toiznt</i> | 'a thousand' |
| Retention of SEY monophthong: | <i>luz</i> | 'louse' |
| Free variation: | <i>toiz ~ tuz</i> | 'ace' |

U. Weinreich (1958:233) continues, in reference to the above speaker: "Her older sister, brought up in a small village about 50 km to the north, has *loiz*. She also uses *tub* and *zuber*, whereas her children have *tub*, but insist on *zoiber*." Weinreich (1958:233) concludes: "The picture is one of typical geographic staggering, not of phonological conditioning."

4.42. StY *oj* diffused phonologically into NEY, but not so into SEY. Phonological strategies/mechanisms were available in NEY, but not in SEY. Compare the development of the following two HA-origin words—StY *toəs* 'mistake' and *rəfuə* 'remedy'—from Tiberian Hebrew times until their modern forms in NEY and SEY, respectively.

²⁵There are arguments for recognition of residual vowel length at one position—the high front unround vowel /i:/ (vs. /i/), versus an alternative analysis based on vowel quality: /i/ vs. /i/; see U. Weinreich 1958a:223. Weinreich notes differences in phonetic length, but not phonemic length.

HA

*tɔ:fu:θ >

*rəfu:ʔ >

PY

*tɔ:əs >

*rəfu:ə >

NEY

/toəs/ → [tojəs]

/rəfuə/ → [rəfujə]

PSY

*tu:əs >

*rəfy:ə >

*rəfi:ə >

*tuuəs >

*tuwəs >

*rəfi:ə >

*rəfijə >

SEY

tuvəs

rəfijə

Both NEY and SEY generally lost vowel length. As argued above, the presence of a *w* glide required recognition of vocalic length. The claim could be made that a single process of intervocalic hardening developed in both NEY and SEY as a way of eliminating anomalous vocalic length. Unlike NEY however, SEY developed neither GS (glide substitution) nor GI (glide insertion) as phonological strategies. Thus, the phonologies of NEY and SEY differ in crucial ways; the surface similarities do not entail structural identity.

Consider the following: Both NEY and SEY have the glide *j* in the word for 'remedy.' NEY shows the results of the application of GI in *rəfu-j-ə*. GI is a general hiatus breaker in NEY, both word-internally, as in *rəfu-j-ə* and *to-j-əs*, and across word boundaries, as in [der id] 'the Jew,' but [di jidn] 'the Jews.' SEY, however, does not contain a process of GI. The presence of glide [j] in SEY *rəfi-j-ə* is not the result of GI. Rather, it may be viewed as part of a general hardening of glides in intervocalic position in SEY, in which *j* < *i* and *v* < *w*, as in, e.g., SEY *rəfijə*, *tuvəs*. SEY glide hardening must have occurred after the Proto-Southern Yiddish (PSY) fronting of PY **u*:₅₂ > PSY **y*: > **i*:, since the form obtained is SEY *rəfijə*, not ***rəfuvə* or ***rəfivə*. This provides evidence that the glide which hardened was not back/round.

Consideration of glide hardening and GI also sheds light on the chronology of length loss in NEY and SEY. In SEY, vowel length must have been present at least until (SEY) glide hardening appeared, since PSY **tuwəs* > SEY *tuvəs*. On the other hand, in NEY, where

loss of vowel length occurred very early, the second *u* of PY **rəfuuə* was lost before the gliding to *w* and (NEY) hardening to *v* could occur. Later, the general NEY rule of GI inserts *j* in both *rəfu-j-ə* and *to-j-əs*. Thus, based on a comparison of NEY and SEY, it may be concluded that SEY did not develop a process of GI.

SEY also lacked the process of GS, as found in NEY. SEY speakers show: (1.) retention of an older dialectal monophthong, e.g., *luz* 'louse,' (2.) lexical diffusion of a StY diphthong — e.g., *ojz* 'house,' *tojznt* 'a thousand,' or (3) free variation: *tojz* ~ *tuz* 'ace.' SEY does not show systematic areas of phonological compromise forms, such as were found in the NEY *uj* area.

Consider the additional variants, listed by Weinreich (1958:244), of SEY *tuvəs* : *tus*, *tuus*, *tues*. The variant *tus* may be derived via the following processes, in chronological order: (1.) syncope of the unstressed vowel *ə* (Weinreich uses symbol <e>), followed by (2) regular SEY loss of vowel length; thus: **tuues* > *tuus* > *tus*. Loss of the *ə* (<e>) meant that the second *u* was no longer intervocalic, and thus could not undergo gliding and hardening > *w* > *v*. The variant *tuus*—like regionally-limited *huus* (vowel 54) in parts of NEY—may be seen as an unstable residue of length in a lengthless system. The remaining variant—*tues*—presents a problem since the unstressed vowel <e> remains, but no hardened glide ([v]) is found. However, the general development in SEY of this vowel (vowel 12) followed by another vowel is *-uv-*, as in SEY *tuvəs*, *anuvə* (< HA *hānɔ:ʔɔ:*) 'pleasure' (cf. StY *hanoə*), *nuvnt* < PY **nɔ:ənt* 'near,' etc.

5. Conclusions. In the present paper I have attempted to shed light on the *where* and *how* of the diffusion of StY *oj₅₄* in EY dialects. From the late 19th century and well into the 20th century there has been a significant trend toward standardization across the EY speech territory. Since Vilne/StY *oj₅₄* represented a prestige pronunciation, somewhat similar diffusion patterns might have been expected across the EY dialects—which, based on evidence related to the diffusion of other StY features, were all presumably open to the ongoing sociolinguistic pressures of dialect-leveling and standardization.

While the diffusion of StY *oj₅₄* may thus be viewed as sociolinguistic in origin, the nature of the incorporation of this feature has varied significantly according to structural factors in the borrowing dialects. A relationship was shown to exist between vowel length and the glide

w; those Yiddish dialects which have lost distinctive vowel length have generally lost the glide *w* as well. For the most part, StY *oj*₅₄ has diffused into areas where distinctive vowel length had been lost, or was on the verge of being lost. Thus, the primary areas of diffusion were, as discussed in the present paper, to be found in the NEY and SEY areas. Consistent with our length-related analysis is, furthermore, the spread of StY *oj*₅₄ into former CY—now NEY—territory noted by Weinreich (1958:262), as well as its non-diffusion into Courland Yiddish, which has preserved a back glide: *au*₅₄. Samogitia, which is geographically intermediate between Courland Yiddish and general lengthless varieties of NEY, preserves a back glide as well: *ou*₅₄.

Both NEY and SEY lost distinctive vowel length, thus favoring the borrowing of StY *oj*₅₄. In NEY, the development of two possibly independent processes—Glide Substitution and Glide Insertion—provided the phonological mechanisms which made possible the phonological diffusion of StY *oj*. In SEY the diffusion of StY *oj*₅₄ was lexical, precisely because SEY lacked the phonological mechanisms.

REFERENCES

- Birnbaum, Shloyme. 1934. Di historye fun di alte *u*-klangen in yidish. *Yivo bleter* 6:25-60.
- Herzog, Marvin I. 1965. *The Yiddish language in Northern Poland: its geography and history*. The Hague: Mouton. (Also Part III of the *International Journal of American Linguistics*, Vol. 31, No. 2.)
- Hoge, Kerstin. 1991. From 'Anti-Daytshmerish' to the 'Hidden Standard': The dynamics of the Yiddish Standardization Process. Unpublished master's thesis. The Ohio State University.
- Jacobs, Neil G. 1990. *Economy in Yiddish vocalism: A study in the interplay of Hebrew and non-Hebrew components*. Wiesbaden: Harrassowitz. (Revised version of unpublished 1984 Columbia University dissertation *Economy in Yiddish vocalism: The case of Central Yiddish*. Ann Arbor: University Microfilms International.)
- _____. 1991a. Comment on Paul Wexler's: "Yiddish—the fifteenth Slavic language." *International Journal of the Sociology of Language*, 175-181.
- _____. 1991b. "Glide substitution in Northeastern Yiddish.." Unpublished paper read at Linguistic Society of America Annual Meeting, January 4, 1991, Chicago.
- _____. (forthcoming). "Structure, standardization, and diglossia: The case of Courland Yiddish." To appear in: Dagmar C. G. Lorenz and Gabriele Weinberger (eds.), *Insiders and Outsiders. German-Jewish, Yiddish and German Literature and Culture in Contact*.

- Jakobson, Roman. 1938 [1962]. "Sur la théorie des affinités phonologiques entre des langues." In: *Selected writings*, Vol. 1, 234-246. Mouton: The Hague. Reprinted from *Actes du Quatrième Congrès Internationale de Linguistes*, 48-59. Copenhagen: Einar Munksgaard.
- Kalmanovitch, Zelig. 1926. "Der yidisher dialekt in Kurland." *Filologiske skrifter fun Yivo* 1:161-188. Vilne.
- Katz, Dovid. 1982. Explorations in the history of the Semitic component in Yiddish. Unpublished doctoral dissertation. University of London.
- LCAAJ. 1992. = *Language and culture atlas of Ashkenazic Jewry*, I. Marvin I. Herzog, Uriel Weinreich, and Vera Baviskar, eds. Tübingen: Niemeyer.
- Schaechter, Mordkhe. 1969. "The 'hidden standard': A study of competing influences in standardization." In: M. Herzog et al, *The Field of Yiddish* 3:285-304. The Hague: Mouton.
- Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Veynger, Mordkhe. 1926, [1927-] 1928. "Vegn yidishe dialekt." *Tsaytskrift*. Oysgabe funem institut far vaysruslendisher kultur, I:181-208; II-III:613-652.
- Vilenkin, Leyzer. 1931. *Yidisher shprakhates fun sovetsfarband*. Minsk: Vaysrusishe Akademie.
- Weinreich, Max. 1923. *Shtaplen. Fir etyudn tsu der yidisher shprakh-visnshaft un literatur-geshikhte*. Berlin: Vostok (Wostok).
- _____. 1960. "Di sistem yidishe kadmen-vokaln." *Yidishe shprakh* 20:65-71.
- _____. 1980. *History of the Yiddish language*. Chicago: University of Chicago Press. [Partial translation of M. Weinreich (Vaynraykh), *Geshikhte fun der yidisher shprakh. Bagrifn, faktn, metodn*. 1973. New York: YIVO Institute for Jewish Research. Translated by Shlomo Noble, with the assistance of Joshua A. Fishman.]
- Weinreich, Uriel. 1952. "Sábesdiker losn in Yiddish: A problem of linguistic affinity." *Word* 8:360-377.
- _____. 1958. "A retrograde sound shift in the guise of a survival: An aspect of Yiddish vowel development." *Miscelánea homenaje a André Martinet, Estructuralismo e historia* 2:221-267. Canarias: Universidad de La Laguna.
- _____. 1958a. "Di klankike struktur fun a podolyer reydenish." In: *Juda A. Jofe-bukh*. New York: YIVO Institute for Jewish Research.
- Wexler, Paul. (1991). "Yiddish—the fifteenth Slavic language." *International Journal of the Sociology of Language*, i-v, 9 -150.

DIALECT ELEMENTS IN SOVIET KOREAN PUBLICATIONS FROM THE 1920S

J. R. P. King
SOAS, University of London

Introduction.¹ Bernard Comrie (1981:8), in his survey of the languages of the Soviet Union, gives two criteria for considering a language as a language of the USSR. It must be autochthonous to Soviet territory, i.e., "...spoken by a population group all or most of whose members were born in the USSR (or territory which is now part of the USSR)," and it "...must not be the official language of a foreign state." Comrie thus excludes Korean and a number of other languages spoken by sizeable groups within the USSR.

The only other mention Comrie gives to Korean is the statement (1981:39): "...Japanese and Korean are not languages of the USSR—although there are 357,507 Koreans in the USSR, mainly in Central Asia, where most of them arrived as refugees from the Japanese occupation of Korea—..."

In fact, there are now close to half a million Koreans in the USSR, and Koreans began emigrating to the Russian Far East in the late 1860s. This emigration frontier was well established by the time Japan annexed Korea in 1910, and Japanese oppression simply increased, rather than initiated, the flow of Koreans to the Russian Far East. The bulk of the Soviet Koreans moved to Central Asia in the fall of 1937, not as refugees, but as victims of Stalin's policy to transfer forcibly all Koreans from the Soviet Far East (see Kho 1987).

Korean is, among other things, a language of the USSR. My fieldwork among Soviet Koreans has shown that their language represents a continuum of northeastern, i.e., Hamkyeng² Province, dialects. This

¹ I am grateful to S. E. Martin for helpful comments on an earlier version of this paper, and to the International Research Exchanges Board (IREX), the Fulbright Commission, and Harvard University's Korea Institute for supporting research trips to the former USSR in the period 1989-1990.

² I transliterate Seoul Standard Korean (henceforth "SS") according to the Yale Romanization, for which see Martin, Lee and Chang (1968, henceforth "MLC"). For the *Hankul*, or Korean script of the Soviet Korean materials and Middle Korean (henceforth "MK"), I use the same system, but with the following modifications: the "alay a" is /ʌ/, the sign represented in SS by /u/ is ㅏ, SS /o/ = /o/, SS /wu/ = /u/, one dot to the left (j/) = high tone, two dots to the left (j:) = rising tone, and low tone is unmarked. I sometimes use the period to make explicit

dialect area in general is poorly understood within Korean dialectology, but the rare northernmost variety, known as 'Yukcin or 'Yuk.up dialect, is now recognized as a highly archaic "dialect island" of great value to Korean linguistics (see King 1987a and forthcoming-c, Zhao and Xuan 1986). I have tried to show elsewhere that pre-Soviet Russian-language materials on Korean, whether written by missionaries, amateurs, or military bodies, contain a wealth of dialect data closely related to the forms used by Soviet Koreans today (see King 1989a, 1989b, 1991).

In this paper, I examine another set of written materials: the Korean handbooks, textbooks and primers published in the Soviet Far East in the 1920s³. The only pre-Soviet Korean-language publications from the Russian Far East of which I am aware are a few newspapers published intermittently in Vladivostok in the years 1908-1917, a religious journal (the *Tayhanin Cyengkyopo* "Bulletin of the Korean Orthodox Faith," published in Chita from 1911-1913), and a handful of other religious—mainly Christian—materials.

The Language of Soviet Korean textbooks, 1923-1929. The Far Eastern period in Soviet Korean history extends from 1923, the year Soviet power was established in the Far East, until 1937, when the Koreans were forcibly resettled to Central Asia. Korean publications from this period have never been studied, although they contain many northeastern dialect elements, especially in their lexicon. While the Korean of these publications is broadly uniform, I distinguish two sub-periods on the basis of trends in Soviet Korean orthography. The first period begins with Soviet power in 1923 and extends until 1930, when O Changhwan established a new standard grammar and orthography with his *Ko.lye Muncen* [Korean Grammar]. The materials examined in this paper are from publications predating O Changhwan's otherwise potentially ambiguous syllable boundaries. See Martin (1993, p. 8): "(1) The dot is used to indicate the "zero" (vowel) beginning of a syllable or other ambiguous situations when the preceding Hankul syllable ends in a **patchim** (final) consonant: **mek.e**, **mek.ko**, **mek.hinta**. Of course, when the boundary is shown by a space or hyphen, the dot is unnecessary: **Puk Han** "North Korea," **kak-kak** "each separately." Korean words written in CAPS represent words written in Chinese characters in the cited texts.

³ I thank Sasha Vovin, formerly of the Institute of Oriental Studies, Leningrad Branch, and Igor' Bol'shakov, formerly of the Korea Section, Leningrad Public Library, for their help in locating these rare materials.

grammar⁴. They are as follows:

- I Ph.Ph., O Sang.il and O Changhwan. 1929. *Say Tokpon (Calani-iy)*. Tyey 2 Kwen. [New Textbook (For Adults). Vol. II.] Xabarovsk.
- Ko.lye I.min Ci.nam. *Hapalopsikhi Sincinsi.khi Ku.yek kwa Kku.h.tta.h.kinsi.khi Ku.yek-iy Hyengphyen kwa isa hanin Pangpep ey tay haye*. 1929. = Čto nužno znat' Korejcu pereselencu. Opisanie Kur-darginskogo i Sindinskogo rajonov v Xabarovskom okruge i pravila pereselenija [Korean Immigration Guide...]. Izdanie Xabarovskoj Pereselenčeskoj Partii. Tiraž 3000. 59 pp. (= "KI29")
- Kongchyeŋ-hoy Haysam-hyen Kanpu wa Wentong Kanpu Palhayng. 1926. *Ppioneyhi Citoca. Tokpon* [Pioneer Leader. A Reader.]. 70 pp. Tiraž: 500. (= "PC26")
- Ni, P., Osanir and Ochanxvan. 1929. *Novyj bukvar' dlja vzroslyx*. Čast' I: Izdanie vtoroe [New Primer for Adults. Part 1, 2nd ed.]. Xabarovsk: Knižnoe delo = Pp. Ph. Ni, O Sang il and O Changhwan. *Say Tokpon (Calan.i-iy)*. Cey-ilkwen. Cey-i.phan. 40 pp. Tiraž: 28000. Čast' II = 99 pp., Tiraž: 28000. (= "NB29")
- Osanir and Očanxvan. 1929. *Metodičeskaja zapiska k bukvarju "Novaja škola"* (na Korejsk. jaz.). = *Tokpon "Sayhakkyo" Kyosu-se* [Teacher's Notes for the Primer "New School"]. Kyosa-yong. Xabarovsk: Knižnoe Delo. (= "Met29")
- P. Ni, T. Ogaj, N. Očanxvan, I. Oseled'ko, S. Txaj, Ja. Tsoj and S. Jugaj. 1929. *Novaja škola: pervaja kniga korejskogo škol'nika* = Li Pyengkuk, O Sangil, O Changhwan, O.ssey.llyeyti.kko I., Yukai Ss. A., Ccoi Ya., Thay Insu. *Say Hakkyo. Tokpon. Tyey-ilkwen, tyey-ilphyen* [New School. A reader. Vol. 1, part 1]. 80 pp. Xabarovsk. (= SH29)
- Say Hakkyo. Tyey-sa.kwen* [New School. Vol. 4]. 1929. = *Novaja škola. Rabočaja kniga dlja 4-go goda obučenija*. 216 pp. Tiraž: 4000. Wentong Kyoyuk-pu Kwahak Pangpep Hoyiy Inka. Xabarovsk: Knižnoe Delo. (= SH29)
- Lossiya Kongsan-tang Yenhay-to Kanpu Nyeca-pu (ed.). 1925. *Nyeca Tayphyo (Caylyo Sucip)* [Woman Representative (A Collection of Materials)]. 279 pp. Vladivostok. (= "NT25")
- Nam Manchun ("Li Congil trans."). 1926. *Appak Patrŋn Ko.lye (Ilpon Chimlyakcuŋy Kŋypan-ha ey issnŋn Ko.lye)* [Oppressed Korea (Korea under the Shackles of Japanese Aggression)]. = Ugnetennaja Koreja. 35 pp. Tiraž 10000. Vladivostok. (= "AP26")
- Pulk.in Ai: Lo.lyek Hakkyo-yong Ko.lye Say Tokpon* [Red Child. New Korean Reader for Use in Worker's Schools]. Yenhay-to Kyoyuk-pu Phyenchan. 1924. 51 pp. Tiraž: 5000. (= "PA24")
- Pulk.in Ai: Lo.lyek Hakkyo-yong Say Tokpon. Tyey-1 kwen, tyey-2 phyen* [Red Child... Vol. 1, part 2]. = *Pervaja Kniga dlja Čtenija*. 1924. 133 pp. Tiraž: 5000. (= "PA24")

⁴ Pre-1930 Soviet Korean orthography contains interesting experimental elements which deserve further study. O Changhwan's grammar is not known in the Republic of Korea. See King (1991-c).

- Pulk:in Ai*: Say Tokpon. Tyey-il-kwen Tyey-il-phyen [Red Child... Vol. 1, part 1]. 1926. *Krasnoe ditja: Korejskij bukvar'*. Xabarovsk-Vladivostok: Tose Cusik Hoysa. 64 pp. Tiraž: 5000. (= "PA26")
- Pulk:in Ai*: Say Tokpon. Tyey-il-kwen Tyey-il-phyen [Red Child. Vol. 1, part 1. (2nd ed.)]. 1926. *Krasnoe Ditja: Korejskij bukvar'*. Izd. 2-e. Xabarovsk-Vladivostok: Tose Cusik Hoysa. 64 pp. (= "PA26")
- Pulk:in Ai*: Lo.lyek Hakkyo-yong Say Tokpon [Red Child...]. Yen hay-to Kyoyuk-pu Phyenchan. 138 pp. 1927? Haysam.wi (Vladivostok).
- Pulk:in Ai*: Say Tokpon. Tyey-samkwen [Red Child. Vol. 3]. 1927. 276 pp. Tiraž: 5000. Xabarovsk-Vladivostok. (= "PA27")
- Pulk:in Ai*: Say Tokpon. Tyey-sa.kwen [Red Child. Vol. 4]. 1927. = *Krasnoe Ditja. Kniga dlja čtenija na korejskom jazyke*. God IV. 272 pp. Tiraž: 3000.
- Sipwel Hyekmyeng Sipcu-nyen Wentong Kiynyem Cunpi Wiwen-hoy (ed.) 1927. *Sipwel Hyekmyeng Sipcu-nyen kwa Ssopeythi Ko.lye Mincok* [The 10th Anniversary of the October Revolution and the Soviet Koreans] = Desjatiletie Oktjabr'skoj Revoljucii i sovetskoe korejskoe naselenie. 109 pp. Tiraž: 3000. (= "PA27")
- Wentong Pyenkang Kyoyuk-pu Penyek-kuk Yeksul. 1929. *Sengin Munmayng Thoychi-so Kanglyeng* [Program for the Eradication of Adult Illiteracy] = *Programma dlja zanjatii so vzroslymi v školax gramoty*. 16 pp. Tiraž: 1000. (= "Pro29")
- Yen hay-to Hakmu.kukto Cyengchi Munhwa-pu Phyenchan. 1925. *Musik ul eps.i han:in Calan.i-iy Tokpon* [An Adult Reader to do away with Ignorance]. 49 pp. Tiraž: 15000. Haysam.wi = Vladivostok. (= "M25")

Dialect Elements in Soviet Korean Publications: 1923-1929.

A. Phonetics and Phonology. In general, it is difficult to draw conclusions from materials written in *Hankul* about dialect phonetics and phonology. In some cases, the Korean script is incapable of reflecting certain features. For example, all Hamkyeng dialects have distinctive pitch-accent and most of them have a flap pronunciation [r] for the liquid /l/ in positions where SS has a palatalized [l']. *Hankul* is not normally equipped to represent such features.

In some cases, it is tempting to see ancient distinctions reflected in certain spellings, but one can never be certain whether the feature in question stems from conservatism in orthography or is a true dialect feature.

Consonantism. With respect to consonantism, the northernmost ¹Yukcin dialects are unique in preserving the syllables /sya, syo, sye, syu/ vs. /sa, so, se, su/, and /t(h)ya, t(h)yo, t(h)ye, t(h)yu/ vs. /c(h)ya,

c(h)yo, c(h)ye, c(h)yu/ vs. /c(h)a, c(h)o, c(h)e, c(h)u/, and Soviet Korean publications from the 1920s also write these distinctions.

The traditional Korean orthography, however, preserved these distinctions long after they had died out in the central dialects, and it is risky to use such spellings in Soviet publications as evidence for the feature.

A related feature of the ¹Yukcin dialect is its preservation of word-initial /n/ before /i, y-/, now lost in SS. Again, the Soviet Korean publications, both pre- and post-1930, write this /n/, but traditional orthography elsewhere did so as well. Some examples from the texts are: /ni, ni.ppal/ "tooth" (PA, PC26) = SS /i, ippal/ < MK /'ni/; /nima/ "forehead" (PA24, SH29, though /ima/ also occurs) = SS /ima/ < MK /ni'mah/; /nyekyesta, nyek.ye, nyekinin/ "consider, deem" (PA) = SS /yeki-/ < MK /nye'ki-/; /nyemlye/ "worry" (PA27, KI29) = SS /yemlye/ < Sino-Korean (henceforth "SK") /nyemlye/; /nyeysnal puthe/ "since olden times" (PA27, but also /yeysnal/ in SH29) = SS /yēys-nal/ < MK /:nyeysnal/; /nyenlyeng/ "age" (PA27) = SS /yenlyeng/ < SK /nyenlyeng/; /nic-/ "forget" (PA27, Pro29, Met29, often also /ic-/) = SS /ic-/ < MK /nic-/; /niese/ "subsequently" (PA27B) = SS /iese/ < MK /nize sye/. The Soviet Korean form is half-Hamkyeng, half-Central in that SS has /iese/ from MK /:nis-/ "connect" > Hamkyeng /[n]is.e/, SS /i[y]e/; /ni-/ "carry on the head" (PA27B) = SS /i-/ (no MK); /niłkhi-/ "bring about" (NT25, SH27) = SS /ilukhi-/ < MK /niłi'thi-/; /nimkun/ "sovereign, lord" (PC26) = SS /imkun/ < MK /:nim'kim, :nim'kum/; /niłi-/ "reach" (Pro29) = SS /ilu-/ < MK /ni'li-/; /niłk-/ "read" (Met29) = SS /ilk-/ < MK /niłk-/; /nyelim/ "summer" (KI29, NB29, SH29) = SS /yelum/ < MK /nye'lim/; /nyeh-/ "insert" (KI29) = SS /neh-/ < MK /nyeh-/ (the SS development is irregular); /niph/ "leaf" (KI29) = SS /iph/ < MK /'niph/; /nip-/ "wear" (NB29) = SS /ip-/ < MK /nip-/; /nił.e na/ "get up" (NB29) = SS /il.e na-/ < MK /ni'lena-/; /niłku-/ "bring about; achieve" = SS /ilwu-/ < MK /:nił.[G]wi-/; /nithancil/ "peat coal; peat" (KI29) = SS /i.than/ < SK /ni.than/ "peat," etc.

In a few cases, the Soviet Korean materials show initial /ni/ or /ny-/ for words which lacked the /n/ in MK. These must be secondary developments: /nikki/ "moss" (PA27, KI29, but also /ikki/) = SS /ikki/ < MK /isk/. KPC has /nikki/ for Cennam, but not Hamkyeng, and KTK does not list /nikki/ for N. Hamkyeng; /nyemil ttay/ "when they

ripen" (KI29) = SS /yemul-/ < MK /ye'mil-/; /nisak/ "ear of grain" (NB29) = SS /isak/ < MK /isak/.

Lack of /p/-lenition. One feature of the consonantism in the Soviet materials which must be closely related to the Hamkyeng dialect background of Soviet Koreans is their largely consistent preservation of unlenited /p/ in "p-irregular" verbs (treated as "w-bases" in MLC). E.g.: /tulyepin/ "frightening" = SS /twulyep-, twulyewun/; /sanap.in/ "violent, vicious" = SS /sānap-, sanawun/; /musyep.in/ "scary" = SS /musep-, musewun/; /CA.YUlop.in/ "free" = SS /cayulop-, cayulowun/; /cilkep.inye/ "enjoyable" (M25) = SS /culkep-, culkewumye/; /etup.in/ "dark" = SS /etwup-, etwuwun/; /kop.in/ "pretty" = SS /kōp-, kowun/; /mukepecesta/ "got heavy" = SS /mukep-, mukewe/; /chip.ecinta/ "gets cold" and /chup.in/ "cold" = SS /chwup-, chwuwe, chwuwun/; /pukkilep.in/ "ashamed" = SS /pukkulep-, pukkulewun/; /tep.in/ "hot" = SS /tep-, tewun/; /ka.kkapin/ "close" = SS /kakkap-, kakkawun/; /pu.lyep.e/ "envies" (PA26) = SS /pulep-, pulewe/; /pankap.ase/ "pleased" (PA27) = SS /pankap-, pankawase/; /chup.i/ "the cold" (PA27) = SS /chwuwi/; /tep.i/ "the heat" (PA27) = SS /tewi/; /tuthepin/ "thick" = SS /twukkep-, twukkewun/; /elyepimilo/ "as it is difficult" = SS /elyep-, elyewum ulo/; /putilepin/ "soft" = SS /putulep-, putulewun/; /nalkhalopimye/ "sharp" = SS /nalkhalop-, nalkhalowumye/; /top.a/ "help" (PA27) = SS /tōp-, towa/; /may.kkilepin/ "slippery" (PA27) = SS /mikkulep-, mikkulewun/; /ka.kkapi/ "close(ly)" (PA27) = SS /kakkai/; /ecilepin/ "dizzy; messy" (PA27) = SS /ecilep-, ecilewun/; /alimtap/ "beautiful" = SS /alumtap-, alumtaw/; /tuthep.i/ "thickness" (PA27, KI29) = SS /twukkey/; /tuthepesninka?/ "Was it thick?" (PA27) = SS /twukkep-, twukkewessnunka/; /swipin/ "easy" (PA27) = SS /swīp-, swiwin/; /koylopim i/ "woes, troubles" (PA27) = SS /koylop-, koylowum/; /usipe se/ "is funny, laughable" (PA27) = SS /wusup-, wusuwe se/; /kapiyepin/ "light in weight" (PA27) = SS /kapyep-, kapyewun/; /kalu nup.il/ "to lie across" (SS seems to have just the causatives /kalo-nwuita, kalo-nwuphita/), but cf. also /nup.e iss-/ "recline" (PA27) = SS /nwup-, nwuwe/; /ttikepin/ "warm" (PA27) = SS /ttukep-, ttukewun/; /haylop.in/ "harmful" (PA27) = SS /hāylop-, haylowun/; /kup.e/ "roast" (PA27) = SS /kwūp-, kwuwe/; /aychelep.in/ "pathetic, touching" (PA27B) = SS /aychelop-, aychelowun/; /kekcyengsilepin/ "worrisome" = SS /kekcyeng-sulep-, kekcyeng-sulewun/, etc.

Occasionally, though, the Soviet materials write an unlenited form, usually only in the verb "help": /toim/ "help" (M25), /toatalna/ "(asked for) help" (PA). The form /swiupkey/ "easily" (PC26) = SS /swiwun, swipkey/ suggests a reanalysis based on a lenited /swiwun/ < */swip-in/.

Velar Palatalization. Hamkyeng dialects other than the northernmost ¹Yukcin varieties often palatalize /ki, kyV/ to /ci, cyV/ (the latter simplifying to /cV/). There are a few such examples from the Soviet materials: /cipninta/ "fixes, repairs (a roof)" (PA24, SH30) = SS /k̄ip-/ "mend, darn (usually clothes)" < MK /:kip-/; /citong/ "pillar" = SS /kitwung/ < MK /kitong/; /cis/ "share, portion" = SS /kis/ < MK /'kic/; /pay.s.cisilk/ "edge/border of stomach" (PA27) = SS /kisulk, kisulak/ < MK /ki'silk/, etc.

In the case of /cilsam/ "weaving (by hand)" (PA27) = SS /kilssam/ < MK /cilsam/, SS has undergone a reanalysis or back-formation, just as in /cimchay/ "kimchee (pickled cabbage)" (PA) = SS /kimchi/ < earlier SK /timchay/ "steeped vegetables," etc.

Vowels. The writing of /ɛy/. The Soviet Korean materials write /ɛy/ in a number of environments where SS now has /i/. Again, it is not clear whether this is due to spelling archaism or to genuine dialect influence (or both), for the highly reliable Kazan' materials from 1901-1904 show unambiguously that ¹Yukcin dialect preserved /ɛy/ (> SS /i/). There are hints in the Soviet Korean materials that /ɛy/ was better preserved after coronals than after velars, but the issue is difficult to resolve.

In the case of Velar + /ɛy/, the writing of /kɛy/ = SS /ki/ is probably a spelling archaism, especially in Sino-Korean vocabulary, e.g.: /kɛy/ "flag" = SS /ki/, /pihayngkɛy/ "airplane" = SS /pihayngki/, /chunkɛy/ "spring season" = SS /chwunki/, /kongkɛy/ "air" = SS /kongki/, etc. (all examples from PA).

Examples of Velar + /ɛy/ in native Korean words are: /yekɛy, kekɛy, cekɛy/ "here, there, yonder" = SS /yeki, keki, ceki/ < MK /inge'kɛy, ke'kɛy, tye'kɛy/, /cohɛy/ "paper" = SS /congi/ < MK /cyo'hay/, /silkiy/ "wisdom, wits" (PA, but also /silki/) = SS /sulki/ (no MK), /sayngkɛy/ "appear" (PC26, once in PA27) = SS /sayngki-/ (no MK), /hamkɛy/ "together" (Pro29) = SS /hamkkey/ < MK /ham'pskɛy/, /kɛy/ "size" (KI29) = SS /khi/ < MK /'kɛy/, /ikɛy/ "beat, win" (SH27) = SS /iki/ < MK /i'kɛy-, i'ki-, etc. Note also the unetymological /kiyтали-/ "wait for" = SS /kitali-/ < MK /ki'tal-[G]u-, and the consistent writing of

/hi/ where SS writes /huy/ (= [hiy]) but pronounces /hi/, e.g., /hin/ "white" (PA) = SS /hiy-/ /hisikhisik han/ (PA27) "greyish" = SS /hiyssik/, /yuh/ "game, amusement" (PA, PC26) = SS (SK) /yu.hiy/, /hi.mang/ "hope" (PC26) = SS (SK) /hiymang/, etc. The Soviet Korean materials also usually write /kei/ for SS /keiy/ "nearly; almost all."

Matters are perhaps more hopeful in the case of coronal + /iy/. Our materials write /iy/ in: /tiy/ "belt" (PA) = SS /tti/ < MK /'siy/, /siyaci'peni, siyapeni/ "uncle-in-law, father-in-law (of woman)" (M25) = SS (SK) /si-/ "in-law" < MK (SK) /'siy/, /otiy/ "mulberry" (PA) = SS /oti/ < MK /o'tay/, /teiy, teiykey/ "slowly" = SS /teti, teti-/ < MK /te'tiy(-)/, /kyenti-/ "endure, stand it" = SS /kyenti-/ < MK /kyen'tiy-, kyen'tay-, /tiy'tiy-/ "tread on" (PC26. SH29 has /tiy'tiy-/ = SS /titi-/ < MK /ti'tiy-/ /s'tiy-/ "acquire a characteristic" (SH27) = SS /tti-/ < MK /s'tiy-/ "wear a belt," /matiy/ "joint; counter for songs" (PA, KI29) = SS /mati/ < MK /matay/. An example of /iy/ after a labial: /i.miy/ "already" (NB29) = SS /imi/ < MK /i'miy/, etc.

In one or two cases, Soviet Korean /iy/ looks convincing because of known secondary developments within Hamkyeng dialects. Thus, King (1991) shows that MK /oy, uy/ can develop to Hamkyeng /iy/, and the Soviet materials have /mal'nyu-/ (PC26) and /malliy-/ (KI29) "make dry" = SS /malli-/ < MK /mal'.[G]oy-/ /cantiy/ "lawn, grass" (SH30) = SS /canti/ < earlier /canstoy/. Likewise, umlaut can front /i/ to /iy/ in Hamkyeng dialects: /skil'e tiyli/ "recruit" (NT25, PC26. Also /tiyl'e pachi-/ "offer up" (PA26)) = SS /tli-/ < MK /'ti'li-, etc.

Preservation of /i/ after coronals. I wrote above that cases of written /iy/ were more "hopeful" after coronals because most Hamkyeng dialects preserve MK /i/ after coronals anyway (SS has raised /i/ to /i/ here). Again, the problem here is spelling archaism vs. dialect feature, compounded by dialect mixture, since some South Hamkyeng dialects share the raising of /i/ to /i/ after coronals with SS. Examples: /kyuchik/ "rules" = SS (SK) /kyuchik/ < MK (SK) /kyuchik/, /achim/ "morning" = SS /achim/ < MK /a'cham/, /maschimnay/ "finally" (PA24, but /machimnay/ in PA26) = SS /machimnay/ < MK /ma'cham' may/, /il'eki/ "early" = SS /ilcciki/ < MK /il'(c)ek/, /kecismal/ "lie" = SS /kēcis-mal/ < MK /'ke'cis-, /ecilepin/ "dizzy; messy" (PA27) = SS /ecilep-/ < MK /'e'cilep-, /kichim/ "cough" (PA27) = SS /kichim/ < MK /ki'chim/, /pucilen ha-/ "diligent" (PA27) = SS /pucilen ha-/ < MK /pi'cilen/, /umeiki-/ "move" (PA27) = SS /wumciki-/ < MK /umei'ki-, /eimsayng/

“animal” (PC26) = SS /cimsung/ < MK /cimsayng, cimsing/, /aɕik/ “(not) yet” (SH27) = SS /acik/ < MK /anʰɕik, anʰcik/, /cɪnpheɭ/ “swamp, bog” (KI29) and /cɪl-/ “muddy” (SH29) = SS /cɪl-/ “muddy” < MK /cɪl-/ , /nameɪs.i/ “more than, upwards of” (NB29) = SS /namcisi/ < MK /namcas hɐ-/ , /aɕimeni/ “aunt” (SH29) = SS /acwumeni/ < MK /aʰcaʰmi/, /chɪlk.holangi/ “striped tiger” (SH29) = SS /chɪlk/ “striped” < MK /ʰchɪlk/, /aɕipani/ “uncle” (SH29) = SS /acay, aɕessi/ < MK /acaʰpa:nim/, /oɕik/ “only” (SH27) = SS /oɕik/ < MK /oʰcik/ , later /oɕik/, etc.

/ɬ/ for SS /wu/. In a few cases, usually after labials, the Soviet materials write /ɬ/ where SS has /wu/. E.g.: /etɪpɪn/ “dark” = SS /etwup-, etwuwun/ < MK /eʰtɪp-, eʰtɪWɪn/, /wensɪngi/ “monkey” (PA27) = SS /wenswungi/ < earlier /wensɪngi/, /tali-/ “to treat” (KI29) = SS /talwu-/ < MK /talʰho-/, /moksim/ “life” (SH27) = SS /mokswum/ < MK /mok:sum/, /kamɪle/ “weather is dry” (PA27, KI29) = SS /kamul-/ < MK /ʰkaʰmal-/, /pata.mɪl/ “ocean, sea” (PA27 = SS /pata/ “sea” + /mul/ “water” < MK /ʰmɪl/, /a.mɪl/ “the Amur” (PA27), /pɪl/ “fire” (rarely) (PA27) = SS /pul/ < MK /pɪl/, /yemɪnin/ “ripening” (KI29) = SS /yemul-/ < MK /yeʰmɪl-/, /ppɪli/ “root” (KI29) = SS /ppuli/ < MK /pulʰwi/, /memɪl-/ “stop, stay” (KI29) = SS /memul-/ < MK /meʰmɪl-/, /phɪlsu.phɪm/ “essentials” (SH27) = SK /phɪlswu.phwum/, /pɪ.the/ “starting from” (SH27) = SS /puthe/ < MK /pɪʰthe/, /tɪmɪn/ “sometimes, occasionally” (PA27) = SS /tumun/ < MK /ʰtɪmɪl-/, etc.

Lack of Umlaut. The lack of umlaut in certain words is another feature more easily ascribed to spelling conservatism than to dialect influence, as most Hamkyeng dialects have undergone umlaut to varying degrees. Examples: /talakki/ “basket with small opening” = SS /talaykki/, /tali-/ “hit, strike” (PA, PC26) = SS /ttayli-/, /kiltongi ɬɪ taliko/ “taking Kiltong-i with him” = SS /teyli-/ < MK /tɐʰli-/, /ca.mi iss-/ “is fun, interesting” = SS /caymi/, /kongcungcep.i/ “somersault” = SS /kongcwung-ceypi/, /kwangi/ “hoe, pick” (PA27) = SS /kwayngi/, /palamkap.i/ “toy pinwheel” = SS /palam-kaypi/, /nali-/ “descend, come down” (PA27, KI29) = SS /nayli-/ < MK /nɐʰli-/, /cokcep.i/ “weasel” (PA27) and /cokcep.i/ (SH29) = SS /cokceypi/, /he.yemchi-/ “to swim” (PA27) = SS /hey.em chi-/ < MK /heʰye/, /kokkwangi/ “hoe, pick” (PC26) = SS /kokkwayngi/, /case.hi/ “in detail” (PC26) = SS (SK) /casey-hi/ < MK /ʰcasyeʰhi/ > SS /casey-hi/ by way of */casyeʰ-hi/, etc.

More interesting are the few cases where this dialect seems to have either undone or avoided undergoing monophthongization of /ai/ > /ay/ and /ei/ > /ey/, a phenomenon identified in King (1991), but the Soviet materials contain few such cases. E.g.: /sil-ɨl naiko/ "produce thread" (PA27) = SS /nāy-/ < MK /:nay-/.

/o/ for SS /wu/. Another feature of only minor importance is the preservation of /o/ in non-first syllables where SS now has /wu/, e.g.: /kocho/ "pepper" = SS /kochwu/ < MK /kochyo/, /tathom/ "fight" = SS /tathwum/ < MK /tʌˈtho-/, /ssahom/ "fight" = SS /ssawum/ < MK /(s)saˈho-/, /so.kkopcil/ "playing house" = SS /so.kkwupcil/, /caco/ "often" = SS /cacwu/ < MK /cʌˈco/, /tomoci/ "(not) at all" = SS /tomwuci/ < earlier /tomoci/, /aco/ "very" = SS /acwu/ < earlier /aco/, /hocho/ "pepper" (M25) = SS /hwuchwu/ < SK /hocho/, /nacong ey⁵/ "in the end" (PC26) = SS /nacwung ey/ < MK /:naycyong/, etc.

The Rule /l/ > Ø/[__+coronal]. I have shown elsewhere that most Hamkyeng/Soviet Korean dialects have a synchronic rule dropping /l/ before coronals, but the Soviet materials from this period show only one example: /noca/ "Let's play!" (SH29) = SS /nōlca/ < /nōl-/.

B. Morpho-Syntax. While orthographic traditions stymie most attempts to glean phonetic and phonological information from the Soviet materials, they are less apt to hide dialect features in the morphology and syntax.

Dialect Features in Nominal Morphology.

/-[k]key/ animate dative-locative (= SS /-eykey/). /mankimi-key nɨn pusi epsta/ "Mankimi has no writing brush" (PA24), /mal kkey cim-ɨl sil.s.ko/ "load luggage on a horse" (PA27), /citoca ɨl key/ "to the leaders" and /nongmin ɨl key/ "to the farmers" (PC26), /lolyekca key/ "to the worker" (NB29), etc.

Vowel-initial suffix forms on vowel-final nouns. In King (1991), I showed how MK word-final /-h/, lost elsewhere in Korean, has survived underlyingly in Hamkyeng as a sort of "ghost consonant" which blocks the usual umlaut caused by affixing of /-i/, and causes these nouns to behave like consonant-final stems with respect to two-shape particles. The Soviet materials contain a few such examples: /khoi/ "nose; net

⁵ Following the practice in Martin (1993, pp. 18-20), I write particles separately. The hyphen is used "...as an unobtrusive way to show the first layer of internal structure of some of the words." (ibid., p. 20).

mesh" = SS /kho [ka]/ < MK /'koh/, /h̥ilo/ "along the top" (also: /u eyse/ "from the top," etc.) = SS /wi lo/ < MK /uh/ (PC26 even writes /h/ once: /u.h̥i lo/ "upwards"), /tw̥ilo/ "to the rear" = SS /tw̥i-lo/ < MK /:twih/, /nai elmanya?/, /ne nail.../ "age" = SS /nai [ka], nai [lul]/ < MK /'nah/, /theilu/ (PA27) "towards a place" = SS /the/ < MK /'theh/, /coi/ "millet" = SS /co/ < MK /coh/, /yelei ... kanta/ "many go" (SH29) = SS /yeles i/ < MK /ye'leh/, etc. In the case of /mulkailo/ "towards the edge of the river" = SS /mul-ka/ < MK /'m̥ils:kaS/, either MK /-s (-z)/ went to /h/, or this is a different problem.

There are also examples of "ghost" /h/ preventing monophthongization in verb stems: /peinta/ "cuts, reaps" = SS /p̥ēy-/ < MK /pe'hi-/ (also /pehi-/ (KI29)), /neelee se/ "gnaw, and..." = SS /n̥ēl-/ < MK /ne'h̥il-/ , /caila/ "measure!" = SS /c̥āy-/ < MK /ca'hi-/ (cf. /'cah/ "ruler"), /ppye h̥il ein̥n̥ h̥is/ "as if cutting through to the bone" (PA27B) = SS /ēy-/ < MK /e'hi-/ , /pui-/ "white, milky" (KI29) < MK /puh̥iy-/.

In some case, the materials write /h/ where SS does not: /kaskahi/ "close(ly)" (PC26) = SS /kakkai/ < MK /kaska'i, kaska'Wi/, /munhi-/ "to destroy, tear down" (PC26) < earlier /munh̥iy-/ (cf. SS /mune-cita/ "collapse"), /sohaci/ "calf" (PA26) = SS /songaci/, /k̥k̥h̥n̥n̥/ "pulling (a load)" (SH29) = SS /k̥k̥l̥-/ < MK /k̥is-/ , etc.

Instrumental /-h̥i]lu/ = SS /-[u]lo/. The Soviet materials have /-h̥i]lu/ everywhere, and once they even write the rarer /-h̥i]llo/: /sip̥:llo/ (KI29) "by/on account of the damp."

Nouns in Final /-i/. In SS, personal names with the diminutive /-i/ drop the /-i/ before adding, e.g., the vocative: /poknam-i/ "Poknam," but /poknam-a!/. In MK, words for animate nouns ending in /-i/ did this as well, and our materials have similar examples, e.g., "goose": nominative /kileki... kanta/, vocative /kilek-a/, topic /kilek-in/, comitative /cek̥-in̥ kilek kwa kh̥in̥ kilek/ "big geese and small geese." Examples with children's names: nominative /mankim̥.i ka/, genitive /mankim̥iy /, etc.

The Plural Morpheme. The Hamkyeng form is /tel/, but our sources usually write /til/. Nonetheless, there are some cases with /tel/: /kangmul tel/ "the streams" (KI29).

Dative /talye/ = SS /tele/. /sunkil.i talye oyyang-il chila hayesta/ "...told Sunkil to clean up the barn" (SH29). From MK /t̥a'lye/ < /t̥a'li-/ "take someone along with you."

Noun + /-cil/. Soviet Korean makes far greater use than SS of the

post-noun /-cil/ "act of doing": /pangmangi-cil ha-/ "to hammer with a cloth fuller," /machi-cil/ "hammering," (M25), /thop-cil-il hako ... tayphay-cil-il hanta/ (PA 1926) "saws and planes," /sol-cil/ "brushing" (SH29), /pay-cil/ "boating," etc.

Dialect Features in Verbal Morphology.

/-emcik ha-, -amcik ha-/ = SS /-um cik ha-/. /mekemcik hata/ "looks good to eat" (PA24) < /mek-/ "eat," /toyemcik ha-/ "looks like it could be" < /toy-/ "become," /poamcik ha-/ "is pleasant to look at" (NB29) < /po-/, etc. The Hamkyeng vocalism preserves MK /-amcik, -emcik/.

Causative morphology. Hamkyeng usually has [-gu-] or [-u-] = SS /-i-, -li-, -hi-/ for the causative morpheme. Examples in /-ku-/: /tolkunta/ "makes turn; spins it" (PA24) = SS /tollu-/; /sokku-/ "to thin out" < *sos-ku- "make spring forth," /kelku-/ "make walk" (AP26) = SS /kellu-/. An example with /-ki-/ = SS /-li-/: /talkye tinta/ "rushes at, attacks" = SS /tallye tinta/.

Causatives in /-u-/: /nayliwesta/ "made come down" < /nayli-/ "descend," /ciwesta/ "defeated" < /ci-/ "lose," /haypich ey palayu-/ "to bleach (it) in the sun" (PA27) < /palay-/ "it fades," /calayu-/ "raise, make grow" (PA27; PC26 has /caliu-/ < /cala-/ "it grows," /hiliu-/ "make cloudy, confuse" (PC26) < /huli-/ "is cloudy," /path-il swiu-/ "let a field rest" (KI29) < /swi-/ "rest," /mul-il oyunin kes/ "irrigation" < SS /ōy-/ "it moves aside" (KI29), etc. Some of these /-u-/ causatives are pleonastic, attached to verbs already causative in meaning: /tiliu-/ "let hang down" (PA27) < /tuli-/ "let hang down," /cim-il puliu-/ "to unload" < /puli-/; /kaliu-/ "cover, hide" (PA27) < /kali-/ "cover it," /coyu-/ "tighten, stretch" (PA27) < /coy-/ "stretch it," etc.

Other causatives (non-Hamkyeng): /cec-il swiie/ "ferment milk" (PA27) < /swi-/ "turns sour," /etuphi-/ "make dark" (PA27) < /etwup-/ "dark," /path-il muk.i-/ "let a field rest" (KI29) = SS /mukhi-/ < /muk-/ "lies unused." The following causatives are spelled in MK-style, as if implying /-l.[G]-/: /sal.ye/ "save somebody's life" (PA27B) < MK /sal.i-/ = /sal.[G]i-/ > SS /salli-/ (SS /sāl-/ "live"), /al.ye cu-/ "let know" (PC26) < MK /al.[G]oy-/ > SS /alli-/ (SS /āl-/ "know"), /tol.ye ka-/ "pass around" (PC26) < MK /tol.[G]o-/ > SS /tollu-/ (SS /tōl-/ "revolve"), /ol.ye/ "make go up" (PC26) < MK /ol.[G]i-/ > SS /olli-/ (SS /olu-/ "ascend"), etc.

Passives. Many Hamkyeng passives tend to have two morphemes,

passive + causative /-wu-/ ("let oneself get VERB-ed") where SS has just one (passive): /caphiul kkataiki/ "a reason to be caught" and /may eykey caphiuko/ "is caught be a hawk" (PA27) vs. SS /caphi-/ < /cap-/; /koyangi eykey hepiueta/ "was scratched by the cat" (PA26) < /hepi-i-wu-/ < /hepi-/ "scratch," /pi ey sich.iuci anhko/ "does not get watered by the rains" (PA27) < /sich-i-wu-/ < /sich-/ "wash" (also /pi ey sichkiun.../ "(plants) watered by the rains" (PA27) and /pi ey nil sich.ie/ "always gets watered by the rain" (PA27)); /hwitulliu-/ "be brandished, played around with" (PA27) < /hwitwul-li-wu-/ < /hwitwulu-/ "brandish;" /malliu-/ "to dry; be dried" (PA27); /pul ey ssaiu-/ "be enveloped in flames" (PA27) < /ssa-i-wu-/ < /ssa-/ "wrap;" /ppuliu-/ "gets splashed, splattered" (PA27) < /ppuli-i-wu-/ < /ppuli-/ "sprinkle;" /kilkhiu-/ "get scratched" (PA27) < /kulk-hi-wu-/ < /kulk-/ "scratch;" /mekhiu-/ "gets eaten" (PA27) < /mek-hi-wu-/ < /mek-/ "eat;" /peli eykey ssoiesil ttay/ "when stung by a bee" (PA27) < /sso-i-wu-/ (with loss of /w/) < /sso-/ "sting;" /kamchuieta/ "was hidden" (PA27.226) and /kungkwel ey kamchiwa issnin kim songaci/ "the Golden Calf hidden in the palace" (PC26) < /kamchwu-i-wu-/ (with loss of /w/) < /kamchwu-/ "hide;" /phum an ey ankiwa se/ "be embraced to her bosom" (PC26) < /an-ki-wu-/ < /an-/ "embrace;" /pal ey... ttiyu-/ "gets stepped on by a foot" < /tti-y-i-wu-/ < /tti-y-/ "tread on," etc.

Again, some passives have MK-type spellings, e.g., /kimul ey kel.ye se.../ "got caught in a net, and..." < MK /kel.i-/ = /kel.ʔ[G]i-/ > SS /kelli-/ (SS /kel-/ "catch on"), /mun i yel.iko/ "the door opened and..." (PA27B) (no MK) = SS /yelli-/ (SS /yel-/ "open"), etc. Others: /hilk i phainin kes/ "the way the earth was dug up" (PA27) < /pha-/ "dig," /tatki-/ "gets closed" (vs. SS /tathi-/ (PA27), /phyeuci mos hako/ "(seeds) don't get spread out" (PA27B) < /phye-/ "spread out," etc.

Final Aspirates. In a number of verb bases which end in aspirates in SS and MK, the Soviet materials have reanalyzed them as ending in /i/: /kiphta/ "is deep" = SS /kiphta/, /nophita, nophikey, nophici, nophitenya, nophiiko/ "is high" = SS /nophita/, /kathita, kathiki/ "is the same" = SS /kathta/, /alko siphina/ "does it want to know?" (PA27) = SS /siphta/, /puthiki/ "attach" (PA27B) = SS /puthta/, etc.

Another verb which has been reanalyzed as ending in /i/ is "plant," though the history is different: /simininya?/ (M25), /siminta, simiko, similila, simiten/ (PA24), /simimyen/ (PA26), /simici, simiki, simiesten, simikey/ (PA27), etc. SS has /sim-/ < MK /simk-, sim'ke/.

Reanalysis of /-n ka po-/ “seems to be” as /-nkap-/ or /-nkapi-/. /nuku-iy sayngil in kapta/ “it seems to be somebody’s birthday” (PA24), /ttenalyerⁿ kes inkapta/ “it looks as if it is about to depart” (PA24), /sencyen wiwen inkapita/ “seems to be a propaganda member” (PA27), /musin yenko issnⁿinkaputa hako.../ “...said: ‘looks like there is some reason’” (PA27), /CHONG i theciesnⁿinkapta/ “the gun seems to have exploded” (PA27).

Adverbative /-ki/. Hamkyeng has /-ki/ = SS adverbative /-key/. The Hamkyeng form is from MK /-kiy, -kay/, but MK also had (more commonly) /-key/, whence SS /-key/: /ppaliⁿki to kanta/ “sure does go fast” = SS /ppalu-key/ “quickly.”

Negation. The striking feature of negation in these materials is the use of the unabbreviated negative particle /ani/, often spelled /an.i/ (sometimes even /anh.i/), usually abbreviated to /an/ in SS < MK /aⁿi/: /ISANG hakey ani pointa/ “do not look strange” (PA27), /chipci ani hakeysta/ “will not be cold” (SH29). The negative copula is spelled /an.ita/ vs. SS /a.ni.ta/.

Most “Long Form” negatives are abbreviated, by dropping the /ha-/, e.g., /ipcek haci an.i/]n ca/ “those who had not registered” (NT25), /thecici ani[ha-] nⁿkes/ “the fact of not bursting” (PA27), /oci ani[]myen/ “if he does not come” (PA27), /poici ani [ha-]nⁿ/ “invisible” (PA27), /haci animyen/ “if you do not do it” (PA27), etc. But occasionally PA27 has /-ci annⁿita/ (SS /-ci anhnunta/) and /-ci anⁿ/ (SS /-ci anh.un/).

Non-final Endings.

/-i]lye/ < MK /-lye/ for SS purposive /-le/. This usage is actually the intensive /-lye/ of SS /-lye [ko] ha-/. E.g.: /kong cap-ilye ta.la kanta/ “runs off to catch a ball,” /kiⁿm may-lye kanta/ “goes to do the weeding,” /nol-lye kasta/ “went to play,” /tⁿalki ta-lye kanta/ “goes to pick strawberries,” /kkoch-ⁿil kkek-ilye ... tola taninta/ “go around to pick flowers,” /namu ha-lye kanta/ “goes to collect firewood,” /sihem chili-lye oassipnita/ “I have come to take the exam” (PA27), /san.yang ha-lye ttenanta/ “departs to go hunting” (PA27).

One sees also /-lya/ and /-lyako/ for the purposive/intensive, e.g., /phul pi.lya kas.sesnola/ “went out to reap grass” (M25), /allyamyen/ “if you wish to know” (NT25), /SIKMUL-ⁿil ku halya tanni-/ “go around in search of food” (PA27), /nollya ko cunpi hanⁿin cha ey/ “just when they were about to have some fun” (PA27B), /toylyamyen/ “if you wish to become,” /hyangsang siki.lya.myen/ “if you wish to improve

it" (PA27B), /cal ha.lya.myen/ "if you wish to do it well" (PA27B), /myen halya ko/ "intending to avoid" (SH27), /HAKKYO lu kalyako MOCA lil ssinintey/ (PA27), /PUNPYEL halya myen/ "if one wishes to distinguish" (PA27), etc.

/haya/ "does/is and...". The Soviet materials usually have the older /haya/ = SS /hay, haye/ < MK /h^hya/, as well as /toyya/ "becomes and..." = SS /tōy, twāy, toye/ < MK /t^hoy^hya/.

Sentence-Final Endings.

/-o, -so/ for the polite-style final ending. The Soviet materials, like Soviet Korean, use /-o, -so/ as a polite-style ending among equals, like Infinitive + /yo/ in SS, whereas /-o, -so/ in SS is the Authoritative or Semi-Formal style (see Martin (1993)), typically used from high to low.

Forms built on the Copula. The Soviet materials contain many examples of final forms built on the copula /-ila/ which are now defunct in SS.

Sentence-final futures: Plain style: /hangpok -in an.i ha.li.la/ "we will not surrender" (M25), /silph.e halila/ (PA24) "you will be sad." /similila/ "will plant" (PA24), /patilila/ "will accept" (PA), /allila/ "you will come to know" (PA27). The same, but Formal style: /tangsin -inhvey lil kaphilita/ "I will repay your kindness" (PA), /pay puhilita/ "you will be full" (SH29), /nay.e culita/ "I will give it to you." (PA27). According to Se Cengmok (1987, 1988), the forms in /-lita/ come from /h^hli-^hngi-ita/, with the MK polite morpheme /-ngi-/, while those in /-lila/ come from /h^hli-ita/, i.e., without the polite morpheme. This analysis helps explain the presence or absence of lenition in the copula (/ita/ vs. /ila/). In Se Congmok's analysis, lenition is blocked by the /-ngi-/.

Forms in /(-ni-)nila/: /KA hanila/ "it is possible," /maynt^hanila/ "are building," /kkoch ila hayesnanila/ "said it was a flower," /cucang hananila/ "advocates," /tongmayng inila/ "is a league" (M25).⁶

Other Endings.

Exclamatory /-nona/: /cal to canona/ "sure does sleep well!" (PA24),

⁶ The first /a/ in the forms with /-nanila/ is a spelling replacement for the old /alay a/ (/A/). Cf. also /kalichinantay lo/ "according to how he teaches" (PC26), /h^hil^hnan/ "flowing" (PC26), /apeci nan/ "as for father" (PC26), /issnani/ (Met29), etc.

/nolay hanuna/ "sings" (PA27) (if these are not elaborate typos for /-kona, -kuna/ = SS /-kwuna/).

Intimate Questions: /molikeysniy?/ "Don't you know?" /elma sik cueskeysniy?/ "How much did you give apiece, do you suppose?" /mues haniy?/ "What are you doing?" /caphwa sangtyem.i issci anni?/ "Isn't there a 5-and-10 store?" This /-niy?/ = SS /-ni?/ is probably < /-ni-[n]-i?/.

Formal Retrospective: /hapteyta/ "(PA27, SH29); Exclamation: /eyki! ...wassokilyel/ (PA27) "My, you've arrived!" Copula in /-ioyta/: /ta phyengting ioyta/ "all are equal" (M25), /loin kwa aioyta/ "... are old people and children" (Met29). For this ending, note that Kim Hyengkyu (1974, p. 388) lists /ha-oita/ for "most of Hamkyeng" and /ha-weyta/ (SS /oy/ is usually pronounced /wey/) as occurring sometimes in both North and South Hamkyeng.

Syntax Proper.

WH-Questions. There are a few traces of the MK system whereby WH-questions took a special set of final question morphemes (e.g., /-ko/ or /-nyo/) different from those in non-WH questions, e.g.: /eccī haye ... chininko?/ "Why do we strike?" (PA24), /mues-il kalichinanyo?/ "What does one teach?" /mues inyo?/ "what is it?" /musin ttis inyo?/ "What does it mean?" /eccī haye ... hananyo?/ "Why do you...?" (M25), /way kilenko hani/ "if you ask why.." (NB29), etc.

Quoted Questions. The system of reported speech in SS allows only Plain style /-nya?/ to be followed by the quotative particle /-ko/, but the Soviet materials have many examples of /-[ni]nka ko/: /khin il ina nasninka ko/ "whether something terrible had happened" (PA27), /tīl.e cukeysninka ko kitali-/ "wait to see whether they would listen..." (PA27), /musin il.inka ko/ "(as if wondering) what was the matter" (SH29), etc.

Case Marking in Relative Clauses. The Soviet materials have a number of interesting examples of downstairs subjects in adnominalized phrases being marked by the possessive particle /y/. This also occurs sometimes in SS, but is usually only recognized for MK: /caki tīl-iy anin tan.e/ "the words they know" (Met29).

C. Lexicon

Without question the greatest utility of Soviet Korean materials written in *Hankul* is in the area of dialect lexicology, and it is here that the

Hamkyeng background of the Soviet Korean authors comes through loudest and clearest. In some cases, SS and Hamkyeng share the same word, and the Hamkyeng reflex differs in minor, but well-established ways. In others, Hamkyeng simply has a completely different word.

Dialect Shapes.

Examples which preserve /-s-/ = SS /-Ø-/ < */z/

/musuntulney/ (PA24) "dandelion" = SS /mintulley/. KTK lists /mwusuntwulley/ and /mwusuntulley/ for most of N. Hamkyeng. The earliest attested form is /miimtilley/ (Tonguy = 1613), suggesting *mizimtilgey.

/keysu/ "domestic goose" (PA27) and /keysani/ "id". (SH29) = SS /kewi/. KTK gives /keysani/ and /keysangi/ for N. Hamkyeng, but neither he nor KPC lists /keyswu/. MK had /ke'yu, 'keyyu/, suggesting *kezyu.

/kkise nay-/ "to pull out" (PA27) = SS /kkül-, kkūs-/ < MK /kīs-, kiz.[G]e/. Cf. also /kīs.e/ "Draw!" and /sengyang-ül kīs.e/ "lights a match" (PA27) = SS /kūs-, kue/, presumably from the same MK root.

/kī pota nāsīn/ "better than that" (PA27) = SS /nās-, naun/ < MK /nas-/.

/pus.in twi ey, pus.e se/ "pour" (PA27) = SS /pūs-, puc/ < MK /pis-, pī.ze/.

/yes.e tilesta/ "eavesdropped" (PA27) = SS /yēs-tut-/ < MK /:yes-, ye.ze/.

/fis.e cu-/ "connect" (PA27) = SS /īs-, ie/.

/ccos.a mek-/ "peck and eat" (SH29) = SS /ccō-/ "peck at; kowtow" < MK /:cos-, coza/.

/kīs.ām/ "weeding" (NB29. Also /kīm/) = SS /kīm, kium/ < MK /ki'zim/.

/meli... ces.āmyense/ "shaking his head" and /no līl ces.e se/ "rowed, and." (PA27) = SS /cēs-, cee/ < MK /ces-, ce.ze/.

Affixation of "Nominative" /-i/.

Soviet Korean dialects typically affix the "nominative" marker /-i/ to all nouns in isolation or in citation form, and sometimes this /-i/ is reanalyzed as part of the base. The Soviet Korean materials have a few such examples:

/peli līl, peli eykey, pel-iy/ "bee (acc., dat., gen.)" = SS /pēl/ < MK /:pel/.

/kolmok.i ey nīn/ "in the street" (PC26) = SS /kolmok, kolmok ey/.

/chacoi to/ "glutinous millet, too" (NB29) = SS /co/ < MK /coh/. Here, the underlying /h/ has blocked fronting of the final /o/.

/thel.i.ka/ "the fur" (SH29) = SS /thel/ < MK /thel/.

In some cases, the incorporation of the /-i/ into the base causes fronting of the preceding vowel:

/pacay/ “fencing wicker” (PA27) and /payc.ay/ (SH29). Cf. SS /pacas-mun/ “twig gate in a bamboo fence.” MK /paco/.

/nokacay namu/ “juniper tree” (PA27) < earlier SK /lo.kacʌ(+ /-i/)/ > SS /noka[n]cwu/ “id.”

/matay/ “yard; threshing area” (PA27). This is SS /matang/. The development must have been MK /math/ > */matah/ > /matang/ > /matang-i/ > /mata[ng]-i/ > /matay/. For the first part of this development, cf. MK /ʼstah/ “earth” > SS /ttang/.

/hephay/ “lungs” (PA) = SS /hepha/ (no MK).

/nunpolay/ “snowstorm” (SH29) = SS /nwun-pola/ (no MK).

In a few other cases, Hamkyeng dialects seem to have avoided, pre-empted or undone monophthongization of /ay/ and /ey/, as if they had reanalyzed the /y/ of such sequences as the “nominative” /-i/. E.g.:

/soma/ “sleeve” (PC26, PA27) = SS /soday/ < /sʰamʌy, ʰsumay/.

/tampa/ “tobacco” (PA27B, KI29) = SS /tampay/.

Preservation of /-k-/ = SS /-Ø-/ < *-g-:

/melku neul/ “wild grape vine” (PA27, KI29). For the first word, cf. SS /melwu/ “wild grapes” < MK /mel.[G]wi/. For the second, cf. SS /nenchwul/, /nengkhwul/. KPC has /newul/ “vine, creeper” for N and S Hamkyeng. PA24 also has /nechul/ “id.”

/kulki/ “a swing” (PA) = SS /kuney/ < MK /kʰil.[G]wi/. KPC gives /kwulki/ for N Hamkyeng, and KTK has /kwulki, kwuli, kwulley, kwulli/. Perhaps < */kʰini-goy/ “swing-thing”.

/molkay/ “sand” (PC26) = SS /molay/ < MK /mol.[G]ay/.

/sulki, sulki tʰil/ “cart, carts” (PC26, KI29, SH29) = SS /swuley/ < MK /sul.[G]wi/.

/elkum/ “ice” (KI29) = SS /elum/ < /ēl-/ “freeze” < MK /:el-, eʰim/. There seems to be some contamination here with another MK verb, /el-/ “to join together, congeal; have intercourse”, related to MK /:el.[G]un/ “adult” > SS /elun/.

/peleci/ (PA24), /pelki/ (SH29) “insect” = SS /pelley/ < MK /pel.[G]ey/.

/namu.kilk.ey .../ “(catch) on a tree stump” (PA27) = SS /kulwu/ < MK /kʰilʰ/.

KTK has /kulki, kulku, kulkey, kulkekci, kulkeli, kulwu/ for N Hamkyeng.

Our materials write /ay.a/ = SS /āy/ for a number of words, e.g.:

/kay.ami/, /ka.ya.mi/ “ant” (PA27) = SS /kāymi/ < MK /kayaʰmi/.

/may.ami/ “cicada” = SS /māymi/ < MK /ʰmʌyaʰmi/.

/pay.ami/ "snake" = SS /pāym/ < MK /ʔpalyam/.

/nay.amsay/ "smell" = SS /nāymsay/ < earlier /nayamsay/.

In a number of words, Hamkyeng has a tense, unaspirated consonant where SS has a plain consonant, or vice versa:

/ttucye/ "turns up (soil with a rake)" = SS /twici-/ < MK /tuʔi-/.

/ssakpalnay/ "laundry done for pay" = SS /saks-ppallay/ < MK /saks/.

/katholi/ "pheasant hen" = SS /kkathwuli/ < earlier /katholi/.

/takcangi/ "ground beetle" = SS /ttakceng-pelley/ < earlier /takcyangpeley, takcyengpelay/. The /-cyang-/ vs. /-cyeng-/ implies */-cyang-/.

/koktayki/ "peak" (PA27) = SS /kkoktayki/ < MK /kok:twi/.

/kkaktakwi/ "striped mosquito" (PA27) = SS /kaktakwi/ (no MK).

/seiley/ "harrow" (KI29) = SS /sseley/. KPC gives /seuley/ as a Ceycwu form.

/temchi/ "straw mat under pack saddle" (SH29) = SS /ttemchi/ (no MK).

/cokci.key/ "small peddler's rack" (PA) = SS /ccok-cikey/.

The shapes of some of the numerals are clearly Hamkyeng: /tasis/ "5" and /yesis/ "6" = SS /tases, yeses/ (which also occur) < MK /ta'sas, ye'sas/, /yetilp/ "8" = SS /yetelp/ < MK /ye'talp/, /ahup/ "9" = SS /ahop/ < MK /a'hop/, /hanna/ "one" (M25) = SS /hana/ < /hanah, hanna/. Another example of Hamkyeng preserving earlier /nn/ > SS /n/: /tanni-/ "walk, go around" (also /tani-/ = SS /tani-/ < MK /taʔni-/.

In at least two examples, the Soviet materials have added /-m/ to a transferentive in /-ta[ka]/:

/katakam katakam/ "step by step" (PA24), and /katakam/ "occasionally" (PA27) = SS /ka-taka/. Cf. also MK /ista'kam/ "in a while".

/yacheyikka il potam phan talita/ "is totally different from ('than') cell work" (PC26) = SS /pota/ "than."

Our materials have /-thuli-/ or /-theli-/ for SS intensive /-ttuli-/: /nem.e theli-/ "to knock over" = SS /neme-ttuli-, /kkaythili-/ "to break, smash it" (PA27) = SS /kkay-ttuli-, and /hekhilethili-/ "to entangle" (M25) = SS /hekhuleci-/ "get entangled" (SS does not have the transitive intensive).

In a few adverbs, our materials write final /h/ for SS /o, wu/: /selh/ "mutually" = SS /selo/ < MK /selh/, /palh/ "straight, directly" (PA27) = SS /palo/ < MK /palh/, /kalnohi-/ "it lies across, sideways" (PA27) = SS /kalo-nohi-/ < MK /kalh/, /koh/ "evenly" (KI29) = SS /kolwu/ <

MK /ko'lo/.

In two intriguing examples, our materials write /ia/ = SS /āy/. King (1991) also reports hearing [ya] = SS /ay/ in Soviet Korean dialects:

/piath.a, piathimyen/ "spit" (PA27) = SS /pāyth-/ < MK /path-/ . MK also had a /:piwas-, :piwath-/ "spit", so the history of this word is problematic.
 /siaksi/ "new bride" (NB29) = SS /sāyksi/. KPC lists this for Hwangnam, but not Hamkyeng. Earlier Korean had /salyaksi/ (Hwae = late 19th c.), /sayaksi/ (Yek Sang = 1690).

The second-syllable /a/ in words like the following appears to be a spelling convention replacing the old /alay a/ (/ʌ/, which should give ʌ/):

/onal/ "today" (PC26, PA27) = SS /onul/ < MK /o'nal/. Our materials also often have /onhəl/ "today" (SH29), a pseudo-etymological spelling based on analogy with /ithul, sahum, nahul/ "1 day, 2 days, 3 days", etc.

/pantasi/ "without fail" (NT25) = SS /pantusi/ < MK /pan'tasi/.

/kalachim/ "teaching" (PC26) = SS /kaluchi-/ < MK /kʌlʌ'chi-/.

/sisalo/ "by itself" (M25) = SS /susulo/ < MK /si'si'lo/, later spelled /sisalo/.

/pitalki/ "pidgeon" (PA27) = SS /pitulki/ < MK /pitur'li/, later /pitulki/ (Tong Uy = 1613), /pitʌlki/ (Pak Cwung = 1677), /pitʌlki/ (Tongmun = 1748).

/hanal/ "heaven" (PC26) = SS /hanul/ < MK /ha'nalh/.

/tutali-/ "rap on, knock on" (PA27; rarely /tu'li-/) = SS /twutuli-/ < MK /tu'li-/.

An MK form */tu'tʌ'li-/ is impossible, as it would have violated vowel harmony, so our form must reflect a pre-modern *spelling* /tutali-/.

It is common for Hamkyeng dialects to raise /Cye/ to /Cey/, e.g.:

/peylak kath.i/ "suddenly" (PA27). Cf. SS /pyelak/ "lightning" < */pye'lak/.

/yeyli-/ "soft" (KI29) = SS /yali-/ , /yeli-/ < MK /ye'li-/.

/himkeyun/ "beyond one's power" (PC26) = SS /him/ "strength" + /kyēp-/ "is beyond one's strength."

/keyca/ "mustard" (PA27) = SS /kyeca/ < earlier /kyecʌ/.

In a few words, our materials have final syllable /o/ = SS /u/ ([i]):

/-lo hayekom/ (marker for causative in causatives) (M25) = SS /-lo hayekum/ < earlier /hʌyekom, hʌyekom/.

/mankhom/ "as much as, to the extent" (PA27) = SS /mankhum/. KTK lists

/mankhom/ for Onseong in the ¹Yukcin region of N Hamkyeng.
 /sokom/ "salt" (PA27) = SS /sokum/ < MK /sokom/.
 /kotonge/ "mackerel" (PA27) = SS /kotunge/ < earlier /kotonge/.

In some cases, our materials have inserted /n/ before /ch/:

/konche po-/ "tries fixing" (PC26) = SS /kochi-/ < MK /ko^hthi-/. The form /konchi-/ is attested in Cheng = 18th c.).
 /kleinchi-/ "quit something" (PC26) and /kleinh.ciela!/ "stop it!" (PA27) = SS /kuchi-, kkuchi-/ < MK /k^hi^hchi-/. /keinchi-/ is attested from the Cheng Tay (= 1728).

In two Sino-Korean words, /n/ has lenited to /l/:

/kay.lyem/ "concept" (PC26) = SS (SK) /kay.nyem/.
 /k^hiy.lyem/ "monument" (SH27) = SS (SK) /ki.nyem/.

In two cases, our materials have second-syllable /u/ < MK /ʌ/:

/han.ul/ "heaven" (PC26) = SS /hanul/ ([hanil]) < MK /ha^hʌlh/. KPC lists /hanwul/ for S Hamkyeng and Kangwen.
 /panul/ "needle" (SH29) = SS /panul/ ([panil]) < MK /pa^hʌnl/. KPS lists /panwul/ for S Hamkyeng, Kangwen and Kyengki.

King (1987a) has shown a widespread rule of /n/-loss before /i, y/ in non-¹Yukcin varieties of Soviet Korean, and our materials have a few examples: /sengyang/ "matches" (PA27) = SS /sengnyang/, /singyangi/ "wolf" (PA27, SH29) = SS /sungnyangi/, /tongyang/ "begging" = SS /tongnyang/.

Some of the kinship terms are clearly Hamkyeng:

/acimi/ "aunt" (M25) = SS /acwumeni/ < /a^hca^hmi/. KTK gives /acimi/ for Sengcin, Kilcwu and Myengchen.
 /apaci/ "father" = SS /apeci/. KTK gives /apaci/ for Myengchen, Kyengwen, Congseng, Hoylyeng and Musan.
 /manulay/ "daughter-in-law" (PA) = SS /manwula/, another good example of final umlaut due to affixation of /-i/.
 /acapi/ and /acaypi/ "uncle" (PA) = SS /acessi/ < MK /aca^hpi/. KTK has /acaypi/ for Sengcin and Kilcwu.
 /hanapeci/ "grandfather" = SS /halapeci/ < MK /hana^hpa:nim/. All KTK's forms show lenition of /n/ to /l/ as in SS, but KPC lists /hanapeci, hanapuci/, etc. for

S Hamkyeng.

Some words show a correspondence Hamkyeng /a/ = SS /e/, often implying MK /ʌ/ or pre-MK */ʌ/:

- /mance/ “first of all” (PC26) or /monce/ “first, first of all” (PA27, SH27) = SS /mence/ < MK /mon(ʰ)cye/ < *manʰcye.
 /monci/ “dust” (PA27, PC26) = SS /menci/ < MK /moncʌy/ < */mancʌy/.
 /posyen/ “Korean-style socks” = SS /pesen/ < MK /posyen/ < */pʌsyen/.
 /phalceng-il ccih-/ “bring one’s sleeves together with arms overlapping within” (PA27) = SS /phal-ccang ul cilu-/ < MK /phalʰtyeng/ < */phalʰtyʌng/.
 /ssakacinta/ “rots” (PA27) = SS /ssek.e cinta/ < MK /sek-/ . The vowel correspondence here could be due to vowel harmonic/isotopic variation rather than to original */ʌ/.
 /ilh.e palyesta/ “lost it” and /paliko/ “throw away” (PA27) = SS /peli-/ < MK /pʌʰli-/.
 /palse/ “already” (PA27) = SS /pelsse/ < MK /pʌʰ(s)sye/.
 /kakay/ “shop, store” (PA) = SS /kākey/ < MK /kakay/ (if the SS /ey/ is not a secondary development from /ay/).
 /ccelnese/ “(too) short, so...” and /ccalneci-/ “gets short” = SS /ccalp-/ < MK /tyeʰʌ-/ < */tyʌʌ-/.

King (1987a, 1991) has shown a wide-spread rule of /ng/-loss in all varieties of Hamkyeng dialects. Our materials have the following examples:

- /mulpaul/ “water-drop” (PA27) = SS /pangul/ “bell; droplet” < MK /pangʰol/.
 /machi/ “hammer” (M25, PA27) = SS /mangchi/.
 /pyeali/ “chick” (PA27) = SS /pyengali/ < earlier /pyengali/.
 /poksua/ “peach” (PA) = SS /pokswunga/ < MK /poksyenghwa/.
 /pue/ “carp” (M25) = SS /punge/ < SK /pu.nge/.
 /komulayngi/ “long-handled prongless rake” (PA) = SS /komulay/ < earlier /komilay, komiley/. Neither KPS nor KTK lists a form with /ng/, and it is possible that it was introduced into our form here by false analogy.
 /kay.aci/ “puppy” (PA) = SS /kangaci/ < MK /kangaʰci, kayyaʰci/.
 /soaci/ “calf” (PA24) = SS /songaci/ < MK /syongaʰci, syoyyaʰci/.
 /tta ka/ “the earth” (PA) = SS /ttang/ < MK /ʰstah/. Here, the SS /ng/ is new.
 /maaci/ “foal, colt” (M25) = SS /mangaci/ < MK /ma.yaʰci/. Here the Hamkyeng form probably introduced the /ng/ and then lost it.

In a large number of examples, Hamkyeng preserves different (usually

older) stem-final obstruents:

/ssikas/ and (once) /ssikac ppilin.../ "seeds" (KI29) = SS /ssias/. KPC has /ssikas/ for S Cenla and /ssikasi/ for S Chwungcheng, but neither he nor KTK has a final /-c/.

/nath/ "face" (M25) = SS /nach/ < MK /nəch/. KPC has /nath/ for S Hamkyeng and Kangwen.

/pis-ɨl/ "debt (acc.)" (PA27B) = SS /pic/ < MK /pit/. The SS form may be from */pic-i/ (with incorporated nominative, or a generalization of the pre-nominative palatalization to the entire paradigm) < */pit-i/. The Soviet form is from the usual merger of /-t/ with /-s/.

/kkoc i/ "flower" (nominative) and /pul kkoc/ "flame (fire-flower)" (PA27, but also /pulkko.chi/ "id".) = SS /kkoch/ "flower" < MK /koc/.

/sich-/ (sometimes /ssis-/) "to wash" (PC26, PA27) = SS /ssis-/ < MK /ʰsis-, ʰsic-/.

/ces-ɨl/ "milk (acc.)" but /cec.i/ (nominative) = SS /cec/ < MK /ʰcyec/.

/cisesta, cisimyense/ "bark" (PA) = SS /cic-/ < MK /eic-/.

It is not clear which dialect assimilated or dissimilated, but KTK has /cis-/ for most of N Hamkyeng. /oyaci[-ka]/ "plum" (PA24) = SS /oyas/ < MK /o.yaɾc i/ (1481). Other MK attestations with final /-s/ are automatic reductions of /c/ before a consonant: /oryas namo/ (Nam 1: 57) and /o'yas nip/ (Kup-kan 6: 29).

/pukk-ɨl totu-/ "heap up the soil around the base of a plant" (PA, KI29) = SS /puk-totwu-/. Cf. Putsillo (1874) *prigrebat' zemlju* [rake up earth (for grain in field)] [putk-i todio]. MK (Cahoy) has /pusto'tol pɔy/, hence *pusk, and KPC lists /pwukki/ for Hamkyeng and Lyangkang.

/kot/ "place" (PA) = SS /kos/ < MK /ʰkot/.

/kyeliph/ "peeled hemp stalk" (PA27) = SS /kyelup/. KTK has /kyeluph/ for Kyengwen.

/poseph/ "plowshare" (KI29) = SS /posup/ < earlier /po'sip/. MLC suggest the etymology /po/ "beam" + SK /sap/ "spade", but this cannot be correct. Cf. Putsillo (1874) *soshnik* [ploughshare] [posyɔppxi]. LCT also lists a /posipsyep/ "side of ploughshare", with /syeph/ = "side". Either Putsillo's form lost the middle syllable via haplology from *posipsyeph, or MK /po'sip/ is from /po'syeph/.

/kaykkichi/ "clean(ly)" (PC26) = SS /kkaykkus ha-, kkaykkusi/. Cf. Xlynovskij (1904) [kiakčiči modxankod, kiakčid xa-] "dirty", and Matveev (1900) "clean" [kaka-zao] (578), "sweet, insipid" [kakadzao] (421). MK had /kaskaci/ "cleanly" (Mong 2 = ?1480s) as the adverb from /kaskashʰ-/, so the Hamkyeng forms preserve early final /-c/ (> SS /-s/). KTK also has /kkaykaci/ for Myengchen.

A number of the Soviet Korean forms preserve /-p-/ = SS /-Ø-/ < *b, another common Hamkyeng feature:

/kupilmyense/ (PA27, SH29) and /kupinin/ "rolling" (SH29) (also /kuil-/ = SS /kwulu-, kwülli-/ < MK /kiul-, kuul-/ < */kibul-/.

/pupuli/ "beak" (PA24, SH29. PA26 has /puuli/) = SS /püli/ < MK /:pu'li/ < */pu'bu'li/.

/akapay namu/ "hawthorn tree" (PA) = SS /akawi-namu/ < MK /aka'oy/ < */akaboy/. /akapay/ is attested in Mulpo (= 18th c.?).

Other Dialect Vocabulary and Expressions.

/olaci anh-/ "does not last long" (PC26), /ola se/ "is a long time since..." (PA27) = SS /olay-/ < MK /o'la-/.

/masi-, mase, masaci-/ "smash; get smashed" (PA, AP26, PA27) = SS /pusu-, puswu-/ < MK /pas-, paza-/. The initial /p-/ /m-/ relation is mysterious.

/pangchi/ "club, mallet" (PA27). Cf. SS /pangmangi/, MK /pangchyu/. If this word is related to SS /mangchi/ "hammer", it is another case of a bizarre /p-/ /m-/ relation.

/momtengi/ "body, frame" (PA) = SS /mom-ttwungi/ < earlier /momstong/. KTK has /momtengeli/ for Myengchen and Hoylyeng, and /momttengi/ for Kilcwu, Myengchen and Musan.

/kak.hi/ "each, each separately" (PA) = SS /kakki/. Neither KPC, KTK or LCT lists a form /kakhi/.

/kili-/ "to do so" (PA) = SS /kuleha-, kulay-/. The Soviet Korean form is an abbreviation for /kuli [ha]-/ "do in such a way."

/sulkalak/ "spoon" (PA) = SS /swut-kalak/ < MK /sul/ + /kalak/. KTK lists /swulkalak/ for Myengchen.

/kilh.im/ "wrong" (PA27) = SS /kulu-/ < MK /kili-/. Our form occurs in the expression /olh.ko kilh.im/ "whether right or wrong", and its /h/ may be from analogy with /olh-/ "correct" < /ol hΛ-/.

/tiley/ "scooper" (PA) = SS /twuley/ < MK /ti'ley/.

/eps.i-/ "to get rid of" (PA27, PC26) = SS /ëps.ay-/. Neither KTK nor KPC lists a similar form, so perhaps ours is abbreviated from MK /:ep'si [hΛ]-/.

/kkwi-/ "dream" (PA27) = SS /kkwu-/ < MK /sku-/. KTK has /kkwi-/ for Congseng and Hoylyeng.

/palim/ "wind" (once in PA27) = SS /palam/ < MK /palam/. KTK has /palum/ for Sengcin and Musan.

/fipsel/ "lips" (PA27) = SS /fipswul/ < MK /fipsi'ul/. KPC has /fipsel/ for Chwungcheng and Cenla provinces, but neither he nor KTK has it for Hamkyeng.

/milu/ "in advance" (PC26, PA27) = SS /mili/ < MK /mi'li/. LCT lists no similar form, but KTK has /milwusi/ for Kyengseng, Onsenseng and Congseng, and KPC has /milusey, milus/ for Ceycwu. I have heard [miru] from Sov. K. speakers.

/cutuli/ "beak" (PA27) = SS /cwutwungi/. KPC has /cwutwuli/ and /cwutwui/ for

N Hamkyeng, and KTK has /cwutwi:, cwuti:/.

/naycong ey nin/ "later; afterward" (PA27; PA27B spells it /laycong/) = SS /nacwung/ < MK /:naycyong/. KTK has /naycwung/ for Hakseng, Kyengwen and Onseng. /mota/ "entirely" (PA, PC26) = SS /motwu/ < MK /mo'to, mo'ta/ < /mot-/ "gather". /cek-/ (occasionally /cyek-/) "be small" (PA) = SS /cāk-/ < MK /:cyak-, :cyek-/. SS /cēk-/ is "few", while /cāk-/ means "small."

/nontilem/ "raised border between two rice fields" (PA) = SS /nontwuleng/ < earlier /nontileng/.

/sonphyek/ "palm of hand" (PA, PC26) = SS /son-ppyek/ < MK /'sons.ʔpyek/ KTK has /sonphyek/ for most parts of N Hamkyeng.

/toyaci/ or /toth/ (PA26) "pig" = SS /twāy-ci/ < MK /toth/, later /tos.aci, toy.aci/.

/pheyki/ "clump, bunch (of greens)" (PA) = SS /phoki/. The Hamkyeng form is due to fronting: /phoyki/ > /phweyki/ > /pheyki/.

/nalsey/ "weather" (PA, SH29) = SS /nal-ssi/. KTK has /nalssey/ for most of N Hamkyeng.

/tungi/ "nest" (PA) = SS /twungci/. KPC lists /twungi/ as a S Hamkyeng form.

/paychay/ "cabbage" (PA, SH29) = SS /paychwu/ < MK /:pɿɿ'chay/, a recent Mandarin loan < báicài. Cf. also /tatu.paychay/ "non-Chinese (long-leafed) cabbage" (PA27, SH29).

/etay/ (sometimes /etey/) "where" (PA) = SS /eti/ < MK /e'tiy, e'tay/.

/paycikey/ "a belly-throw (in wrestling)" (PA) = SS /pay-ciki/.

/ha.li, halhi/ "one day" vs. nominative /halli/ (PA) = SS /halwu/. The spelling with /h/ is pseudo-etymological by analogy with /sa-hul, na-hul/ "3 days, 4 days," etc., but the nominative form is from MK /ha'la, ha'li/.

/hey-/ "to count, calculate" (NT25, PA) = SS /sēy-/ < MK /:hyey-/.

/hey-/ "to swim" (PA). Cf. SS /hey.em/ "swimming", a derived substantive from /hēy-/ which is moribund, and /hey.em chi-/ "to swim". Earlier Korean has only the substantive (/hey.ɨm, heyyum, hey.um, heyyom/) (no MK).

/cay.makcil/ "diving" (PA24) = SS /ca.mayk-cil/. The /ay/ in our form suggests */cyamɿyk/ > */cyaymɿyk/, but more data are needed.

/silh.ye hanta/ "dislikes" (PA) = SS /silh.e hanta/ < MK /silh-/. KTK has /silye ha-/ for Myengchen, Kyengseng and Congseng⁷.

/tolihye/ (PA24), /tol.ohye/ (PA27) /to.lohye/ (NT25) "on the contrary" = SS /tolie/ < MK /tolo'hye, tolɿ(h)hye/. It is normal for Hamkyeng to lose intervocalic /h/ on the surface, so the /h/ in the Soviet forms may be a spelling archaism. See the next example.

/ahay/ (NT25, PC26) and /ahiy/ (SH27) "child" + SS /ai, ay/ < MK /a'hay/. Again, the /-h-/ is probably a spelling archaism.

/kye'il/ "winter" (PA) = SS /kyewul/ < MK /kye'zil(h)/.

/palkwi/ "horse-drawn sleigh" (PA) = SS /palkwu/. Cf. also /ccokpalkwi/ "a sled

⁷ But note that /silye/ for /silh.e/ is also widespread in Seoul dialect (as opposed to Seoul Standard).

- for tobogganing" (SH29). KTK has /palkwi/ for most of N Hamkyeng, and KPC has /ccokpalkwu/ for Cakang and Lyangkang.
- /himk^hei/ "to the utmost of one's strength" (PA) = SS /him-kkes/ < MK /himska^hcang/. KTK has /himkkuth/ for Hakseng and Kyengwen, /simkkuth/ for Kyenghung. SS /-kkes/ is related to SS /-kkaci/ "up to, as far as" < MK /-ska^hcang/, and the Hamkyeng /-kkuth/ is a reanalysis, replacing /-kkes/ with /kkuth/ "end, limit".
- /molo-/ "not to know" (PA27) = SS /molu-/ < MK /mo^hla-/. Neither KPC nor KTK gives this vocalism, but it is the usual reflex for Soviet Korean.
- /kel.eci/ "beggar" (PC26) = SS /kēci/ < MK /kele^hchi/, later /ke^haci, keeci/. KTK has /keleci/ for Kyengseng, Pulyeng, Congseng and Musan.
- /ccol^hni-/ "tighten, strangle" (PC26) = SS /colu-, colla/ < MK /col^h[G]o-/. KTK does not have this reflex, but KPC has it for Hwangnam, Phyengpuk, Kangwen and Kyengki.
- /cokimaksik han/ "each of them a bit small" (PC26) = SS /cokum-ssik/. The final /-k/ in the Hamkyeng form is mysterious, but Soviet Korean also has /hanak-ssi, hanak-ssik/ "one apiece". MK "one" had a final /-h/ (/hanah/) which could conceivably give a /k/, but MK "little bit" (/cyoko^hma/) did not, so this is mysterious. Perhaps these are just anticipations of the /-k/ in /-ssik/, but note that some Hamkyeng dialects have /-ssi/ rather than /-ssik/.
- /che^hli^hm/ "like" (PC26) = SS /chelem/ < earlier /thyelo/. KTK has /chelu/ for Onseng, and /thelu/ for Congseng and Hoylyeng. I have heard /thyelu/ from a Soviet Kyenghung ('Yukcin) speaker.
- /yes^hni^hn/ "put, place" (PC26) = SS /neh-/, an irregular development from MK /nyeh-/ > ¹Yukcin /nyeh-/, non-¹Yukcin Hamyeng /yeh-/. The /-s/ is an orthographic problem (read: /yennin/).
- /kunyeng/ "hole" (PC26, AP26) = SS /kwumeng/ < MK /kumk-/. KTK has /kwunyeng/ for Kyenghung, Congseng and Musan.
- /ppam/ "cheek" (M25) = SS /ppyam/ < MK /^hspam/. Neither KTK nor KPC has this for Hamkyeng
- /sien ha-/ "refreshing" (PA) = SS /siwen ha-/ < MK /siyhwen ha-/. A good example of the tendency for both /-h-/ and /w/ to drop in Hamkyeng.
- /paci ^hil kkeyko/ "put on his trousers and..." (PA27B) = SS /kkwēy-/ "thread through" < MK /pskey-/. Here, SS seems to have innovated its /w/.
- /anlyengi keysio/ "Goodbye". (PA27B) = SS (SK) /annyeng-hi/. A good example of what Ramsey (1978) calls "lambdacism", the rule /nn/ > /ll/ in some Hamkyeng forms (/nl/ is a common early spelling for /ll/).
- /canchay/ "feast, banquet" (M25, PA27B) = SS /canchi/ < MK /^hcanchay/.
- /camkwa se/ "immerses it, steepes it" (PA27) = SS /camku-, camke/ < MK /cam^hk^hal-, camki-/. KTK has /camkwuta/ for Kyengwen and Musan, but our vocalism suggests /camko-/, which looks like a causative.
- /cechi-/ (/cyechi-/ once) "make wet" (PA27) = SS /ceksi-/ (causative of /cec-/ <

MK /ce'ci-/ , later /cechi-/ . The SS reflex is a mystery.

/pupi-/ "rub together" (PA27) = SS /pipi-/ < MK /pu'piy-, pya'pɿy-, pi'pɿy-, pipi'y-i-, spi'piy-/ . KTK has /pwupita/ for all three areas of N Hamkyeng.

/caulteni/ "dozed" (PA27) = SS /cōl-/ < MK /caol-/ . The usual Hamkyeng form is /capul-/ < */caɸol-/ , but KPC has /cawulep-, cawulli-/ for Kangwen.

/cikson/ "pick, hoe" (KI29) (no SS or MK). Perhaps this is related to KPC's /cwuk/ = /hulk-son/ "plasterer's trowel" for Hamnam.

/tuci/ and /tucu/ "wooden grain chest" (NB29) = SS /twicwu/ (no MK). KTK has /twuci/ for most of N Hamkyeng.

/yeyyeng-il yeynta/ "puts roofing on a roof" (NB29) = SS /ieng ul īs-/ , /ieng ul ī-/ < earlier /nyey-, nieng/ (no MK). KTK has both /nieng/ and /yeyyeng/ for N Hamkyeng. An earlier */nyey-/ appears to have raised to /nī-/ to /ī-/ in SS.

/cancali/ "dragon-fly" (SH29) = SS /camcali/ < MK /caŋ'cali/. Neither KTK nor KPC have this reflex for Hamkyeng.

/namu sseki/ "sawing wood" (SH29) = SS /khye-/ < MK /hhye-, hye-/ . KTK has /sse-/ for Kilcwu, and KPC has it for all of Hamkyeng.

/noepko/ "be full of anger, and..." (SH29) = SS /nōyep-/ < MK /:no'hap-/ .

/calu/ "often" (KI29) = SS /cacwu/ < MK /ca'lo, ca'co/. KTK does not list this for N Hamkyeng, but KPC has /calwu/ for N Phyengan and S Hamkyeng.

/phul-seph/ "grass thicket" (PA) = SS /phul-swuph/. KPS has this for S Hamkyeng. The second morph in the SS form is /swuph/ "forest, thicket," whereas our /seph/ must be either SS /seph/ "brushwood" or /seph/ "side, vicinity."

/ko.lye/ (PA) and /ko.lyey/ (SH27) "Korea", /ko.lye mal/ "Korean language" (SH27).

The name of an ancient Korean dynasty (918-1392), /ko.lye salam/ or "people of Koryō" is the self-designation of the Soviet Koreans.

/mul.oy/ "cucumber" (KI29) = SS /oi/.

/mekuli/ "frog" (PA24) = SS /kaykwuli/.

/k[i]nyang/ "continuously; carrying on", e.g., /kacuk.os-il kiyang ipko santako hanta/ "they say they (Eskimos) always/continually wear leather clothes" (PA27). SS /kunnyang/ means "in the same way as before; without doing anything".

/ningkim/ "apple" (PA) = SS (SK) /sa.kwa/. SS /nungkum/ is specifically "crabapple".

/melithay/ "braid" = SS /tayngki/. KTK has /melithay/ for Sengcin, /melttay, melittay/ for other N Hamkyeng regions.

/katayki/ "Korean plow" (PA) = SS /cayngki/.

/poli simin pekim/ "after planting the barley". SS /pekum/ < MK /pekum/ "the second in order" usually appears only adverbially, and never after a modifier.

/cey.a.kim/ "each on its own" < MK /ceyye'kom, cey'kum/. SS has /cey-kakki/.

Neither KTK nor KPC lists this reflex.

/the.path/ "vegetable garden". Cf. SS /the/ "place, site" and /path/ "dry field".

/siksol/ "members of the family" (PA, PC26) = SK /siksol/ "id", but SS uses SK /sikkwu/.

- /changka/ "song". Soviet Korean uses SK /changka/ more than native Korean /nolay/.
- /cinay/ "very; excessively". SS has /acwu, maywu, nemu/, etc. Neither KTK nor KPS lists this word, but I have heard it from Soviet Koreans. Perhaps related to SS /cen-hye/ "entirely, utterly" < MK /cyen-hye/.
- /mul.yek/ "water's edge" (PA, KI29), /pata.yek/ "seashore" (PA27), /kangyek/ "riverside" (PA27), etc. /yek/ is a typical Hamkyeng word for "side of something," and is related to SS /yeph/ < MK /nyekh/.
- /kyeley-capki/ "hide-and-seek". Cf. SS /kyeley/ "brethren" and /cap-/ "catch". SS has /swullay-capki/ for "hide-and-seek".
- /nongsa-them/ "farming grounds/area". Cf. SK /nongsa/ "agriculture" and SS /the/ "place, site", /themuni/ "grounds, basis".
- /cecthong ey cec.i ppule se/ "my udder is full of milk, and so..." Also: /kangmul i ppulesta/ "the river was full of water/overflowing..." Cf. SS /pūt-, pule/ "swells, gets bloated in water; increases" < MK /put-, pu'le/ and /pūs-, pue/ "(part of body) swells". More examples: /etteh.key ppul.enaninka?/ "How do they (the peas) grow/bulge up?" /mak-in ppulesta culesta hamye/ "the membrane expands and contracts alternately" (PA27), /mul i ppil-in ttay/ "times when the water is high" and /inmin i ppulninta/ "population goes up" (KI29).
- /melki/ "waves." Not in SS, but cf. Xlynovskij (1904) [mōl'kī] "wave". KTK lists /melki/ for Sengcin, Kyengseng and Congseng. SS has /mul-kyel/, but the North Korean *Hyentay Cosenmal Sacen* [Dictionary of Contemporary Korean] lists /melki/ as a standard word for "large slow waves kicked up by the wind which persist even after the wind dies down."
- /pi-kochi/ "rain drop", i.e., "rain-flower." Cf. SS /pul-kkoch/ "flame; spark" < "fire-flower."
- /no.hmkam/ "toy". SS has /cangnan-kām/, rarely /nolis-kam/. Cf. SS /nōlum/ "playing, gambling", and /kām/ "material".
- /koki nakk.uki/ "fishing". Cf. SS /nakksi-cil, nakk-/ "id." < MK /naksk-/ [sic]. KTK has /nakkwu-/ "go fishing" for all of N Hamkyeng.
- /phanpakin/ "typical". SS has just the derived noun /phan-paki/ "stereotyped, conventional thing".
- /yesta/ "here you are." Cf. *Razvedchiku...* (1904) [ɛssò], [ɛttà] "Here you are", and Starčevskij (1890) [njotta] "id."
- /ttekcip-ilo myenpo sa.lye kanta/ "I'm going to the bakery to buy bread." (PA) Soviet Korean uses /tek/ "rice-cake" for "bread" now.
- /pal.ka mekninta/ "eats them alive". Cf. SS /pal.ka pes-/ "strips, undresses" (?). Or perhaps this is N Hamkyeng /palku-, palka/ = SS /palu-, palla/ "crack open, husk, shell". Cf. SS /palla mek-/ "to sweet-talk someone out of something".
- /nalsay/ "bird" (PA26). Cf. SS /nal-/ "fly" and /say/ "bird". Neither KPC nor KTK lists this word, but the Kazan' dictionary (RKS, p. 93) records *pütsa* "bird" as: "[sæ], [nar sæ]", with the intriguing note "(pronounced [nar tsæ] in the Maritime Province)".

- /kangtang-iy hanphan ey nin/ "in the center of the classroom" (PA27). KTK has /hanphan/ for most of N Hamkyeng.
- /chaykSANG i nohiko/ "a table was located" (PA27). Hamkyeng dialect seems to use /noh.i-/ (lit.: "be placed") much more often than SS, almost as a substitute for /iss-/ "be located".
- /esilpisi ha-/ "tottering? shivering?" (PA27). MLC has only /esul-esul/.
- /hantay/ "together" (PA27. SH29 has /hantey/). SS has /hamkkey/ < MK /ham'pskiy/, but MK also had /han'tay/ "id". Neither KTK nor KPC lists this, but it is common in Soviet Korean.
- /pomchel i tol.a omyen/ "when spring comes" (PC26, PA27). Soviet Korean uses /tol.a o-/ "come back, return" for the onset of seasons.
- /kangthay/ "moss" (PA27). SS has /ikki/. Neither KTK nor KPC lists this word at all, but MLC lists it as "Dialect".
- /mom-il songkulie/ "curls/coils up its body" (PA27). Not attested elsewhere.
- /tolki/ "a tree-ring" (PA27). Cf. SS /töl-/ "turn, go around".
- /meyng meying/ "sound of a calf" (PA27). The standard SS sound for bovines, regardless of age, is /ummäy/.
- /kkangci/ "dregs" (PA27; PA27B has /kkangchi/). SS has /kenteki/ for "stuff floating in liquids", and /ccikki/ for dregs. KPC gives /kkangci/ as a Kangwen form, and has /kengchi/ as N Hamkyeng for SS /kenteki/, earlier Korean /kenti/.
- /haycapuli/ "sunflower" (PA27, PC26). SS has /hay-palaki/ (lit.: "sun-gazer"). KTK has /haycapwuli/ for most of N Hamkyeng, as well as /haykyapwuli, haykapwuli/. Our form is from palatalization of the /ky/: /hay-kyapul-i/ "the one which leans toward the sun" > /haycyapuli/ > /haycapuli/. Cf. SS /kiul-/ "lean toward", /kyawu-/ (light isotope of /kiwu-/ "move slantwise"), etc. Both KPC and KTK have Hamkyeng /kipul-, cipul-/ = SS /kiul-/.
- /pam tu cyem/ "2 a.m." (PA27) and /achim yel cem/ "10 a.m." (PA27B). SK /cem/ < /cyem/ < /tyem/ "point" where SS uses SK /si/ "time; hour".
- /halhi.kali/ "an acre, hectare" (PA27). Common in pre-Soviet materials as well, this is "a day's plowing": cf. SS /halwu/ < MK /hāl/ + /kāl-/ "plow".
- /kocoli/ "icicle" (PA27) = SS /kotulum/. Earlier Korean has /kos.eüm/ (cf. /eüm/ "ice"). KTK has /kocoli, kotoli, kocolum, kocol, kotulum/.
- /PHIL ha-/ "finish; graduate". SS has SK /phil ha-/ as well, but hardly ever uses it in the spoken language as Soviet Korean does.
- /puntuci/ "cocoon, pupa" (PA27). Cf. SS /penteyki/. KTK has /pwuntwuci/ for most areas of N Hamkyeng.
- /incha/ "now; starting now" (PA27, Pro29). Cf. SS /icey, incey/. KTK has /ica/ for Onsen, and KPC has /finchi/ for N Hamkyeng, but neither has /incha/. The North Korean *Hyentay Cosen-mal Sacen* (1981) recognizes /incha/ as standard, so presumably it is taken from Hamkyeng dialect.
- /tyenhwa hī kimman patasnintey..l/ "I just got a phone call and..." (PA27). Cf. SS /ku man/ "to that extent; so much". This usage is typical of Soviet Korean.

- KPC has /komman/ as N Hamkyeng, Lyangkang for SK /kumpang/ "just now, just when".
- /pulp.e ha-/ "be envious" and /pulp.in maim/ "jealous mind" (PA27, SH29). Cf. SS /pulep-/ < earlier /pīlp-, pulep-/ (no MK). KTK has /pwulpe ha-/ for all of N Hamkyeng.
- /salamsalami/ "people" (PA27). SS normally only reduplicates one-syllable nouns, if at all.
- /yelsi tongan ey/ "for 10 hours". SS would have SK /si.kan/ "hour" for SK /si/.
- /cicnin kichim... kichim il cictaka/ "to cough" (PA27). SS /cic-/ < MK /cic-/ means only "to bark" or "to caw, crow (of a crow)".
- /taltal ha-/ "sweet" (PA27). Unusual, because it is a reduplicated verb base combined with /ha-/ "do". Cf. SS /tal-/ "sweet".
- /chayma cih.i/ "home vegetable gardening" (PA27, KI29) and /chaymatyen/ "vegetable garden" (PC26). Cf. SK /chāyma, chāyma-cen/ "vegetable garden", glossed as "Archaic" by MLC. /cih-/ for SS /cīs-/ "make, build" is typical of the Soviet written materials, but is a spelling convention.
- /ttulp-/ "to pierce through" (PA27. SH29 has /tulp-/). Cf. SS /ttwulh-/ < MK /tulp-, tūlp-, later /stul-, tulp-, stulp-. KTK has /twulpta/ and /ttwulpta/ (also /ttwulka/) for most of N Hamkyeng.
- /nemki-/ "to swallow" (PA27). i.e., "make go over". Cf. SS /nem-/ "go over, cross". KPC has /nengkita/ "swallow" for N Phyengan, and /nengkwuta/ for Hamkyeng and Lyangkang, while KTK has /nemkwuta, nengkwuta, neyngkita/ for N Hamkyeng.
- /nuntakci/ "eyelid" (PA27). KPC gives /nwunttakci/ = SS /nwun-twuteng/ as a S Hamkyeng form, and KTK does not have it for the north.
- /mi.kep-il kolmok ey nay.e palininya? mikepthong ey palininya?/ "rubbish" (PA27. 220). Not attested elsewhere.
- /kkop-/ "stick onto, as in putting arrowheads to a shaft" (PA27.227). KTK has /kkop-/ = SS /kkoc-/ for most of ¹Yukcin and Musan.
- /kayphali/ "dog-sled" (PA27). Cf. SS /kay/ "dog" and /palkwu/ "sled". KTK has a /palki/ and /soy[-]pali/ for sled, but how these all might be related is unclear.
- /il epsta/ "no problem" (PA27). Typical of most northern varieties of Korean.
- /mungki mungki nao-/ "comes forth in puffs/clouds" (PA27). Cf. SS /mungkey-mungkey, mongkay-mongkay/.
- /capi lo/ "by oneself" (PA27B, SH27). Typical of all Soviet Korean, perhaps related to SS (SK) /ca.ki/ < /ca.kiy/, and first attested in Xlynovskij (1904): [tsiāpiro], [tsapi-ro]
- /mecel.i/ "fool" (NT25). KTK has this for most of N Hamkyeng.
- /paywe cu-/ (NT25) and /payhwa cu-/ (Pro29) "to teach". Cf. SS /paywu-/ "learn" and sub-standard English "I'll learn you a thing or two." The /h/ is traditional spelling from MK /pay'ho-/.
- /u.phyen paytal.i/ "the mailman" (PC26). From SK /pāytaal/ "delivery" + /-i/ "person".

- /cangca/ "fence" (PC26). KPC has /cangcay/ for N Hamkyeng.
- /setap/ "laundry" (PC26). Typical Hamkyeng, perhaps related to SK /sēythak/.
- /esi/ "parents" (PC26). Cf. SS /epci/ "parents" < MK /epɛzi/, and MK /e'zi/ "mother; parents". Soviet Korean has [ɔʃidiri] = /esi-tul + i/ "parents".
- /kusipay/ "kind of boat (glossed as /thongmoksen/ "boat from one piece of wood)". (KI29) KPC lists /kusipay/ as Ceycwu for "small row-boat".
- /yaktampay/ "poppy; opium" (KI29). KTK has this for most of N Hamkyeng.
- /poth namu/ "birch". Cf. Putsillo (1874) *berëzy* [birches] [pottsxi], [ponamo] /poschi, posnamo/. Not attested elsewhere.
- /kang kenney/ "across the river" (KI29). Cf. SS /kēnne/ "id". < MK /:ken'ne-, :ken'ni-/. Our form is from the latter's infinitive: */:ken'nie/ > */kennye/ > /kenney/.
- /elimsengi/ (KI29) "frost; ice floes". Cf. SS /elum/ "ice" and /sengey/ "layer of frost; ice drifts (in river)".
- /chachem/ "gradually" (KI29). SS has /chachum, chacum/. KTK lists only /chachum/.
- /niph/ "lake" (given as gloss to SK /hosu/ "lake") (KI29). SS uses SK /hoswu/ for "lake", and SS /nuph/ tends to be "swamp; marshy place".
- /moim/ "meeting" (SH27). Cf. MK /mot-/ "meet, gather". SS has /moim/ < /moita/.
- /kol/ "head" (SH29). This is SK /kol/ "bone; marrow, brain", and is not used for "head" in SS, except as the first member of vulgar compounds.
- /cikakcil/ "sifting (of grain)" (?) (SH29). Not attested.
- /sekma/ "(hand-operated) millstone" (SH29). SS has /māys-tol/. KPC gives a N Hamkyeng /sekmay/, glossed as /yenca-panga/ "millstone worked by ox or horse".
- /meksay/ "food; feed (for animals)" (SH29). SS /mek-say/ is "way of eating; appetite".
- /khongkaktay/ "bean chaff" (SH29). SS has /khong-kkakci/.
- /saykki tali amh.kay/ "a bitch with litter" (SH29). Cf. SS /tal-/ "affix, attach".
- /myenmok ik.in salam/ "somebody one knows well" (SH29). Cf. SK /myenmok/ "features; face" and /ik-/ "get accustomed to".
- /tonglisya/ "general store" (SH29). For SK "same-neighbourhood-company" or "neighbourhood (*tong* + *li*)-store".
- /than/ "snare, trap" (SH29). KPC lists this for Hamkyeng and Lyangkang, and glosses it as "snare, noose". KTK lists it for Hakseng and Onseng.
- /se.phi/ "squirrel" (SH29). From SK /se/ "mouse; rodent" + /phi/ "fur, hide". Common in Soviet Korean dialect, and in pre-Soviet Russian sources.
- /kkweng-iy chakk-il noh-/ "set a pheasant trap" (SH29), /cha.kki/ "snare, trap", and /chakk.ey/ "in a trap" (PA26). Attested in Russian sources, but apparently not elsewhere.

Conclusions. The Korean-script materials published in the Soviet

Far East in the 1920s contain a wealth of data from the Hamkyeng province dialects spoken by most Soviet Koreans. Although devilish problems related to pre-modern orthography render *Hankul* materials from this period almost useless for most phonetic and phonological features of the dialects, other features shine through. Materials like these can shed light on dialect morpho-syntax, but are most useful for dialect lexicography and problems of etymology. They often give us data we might never succeed in eliciting from informants, and provide us with new questions to take to fieldwork in the former USSR.

REFERENCES

- Comrie, Bernard. 1981. *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.
- Kho, Songmoo. 1987. *Koreans in Soviet Central Asia*. Vol. 61 of *Studia Orientalia*, edited by the Finnish Oriental Society. Helsinki.
- Kim, Hyengkyu. 1974. *Hankwuk Pangen Yenkwu* [Research on Korean Dialects]. Seoul: Sewul Tayhakkyo Chwulphan-pu.
- Kim, Pyengcey. 1980. *Pangen Sacen* [Dialect Dictionary]. Phyengyang: Kwahak Paykkwa Sacen Chwulphan-sa.
- Kim, Thay.kyun. 1986. *Hampuk Pangen Sacen* [Dictionary of North Hamkyeng Dialects]. Seoul: Kyengki Tayhakkyo Chwulphan-pu.
- King, J.R.P. 1987a. An Introduction to Soviet Korean. *Language Research*, v. 23, No. 2, June, pp. 233-274.
- _____. 1989a. The Korean Dialect Materials in Matveev's *Reference Book to the City of Vladivostok*. *Language Research* v. 24 No. 2, pp. 281-329.
- _____. 1989b. A 1902 Cyrillic-script Rendition of "Little Red Riding Hood" in North Hamkyeng Dialect. In: Kuno et. al. (eds.) *Harvard Studies in Korean Linguistics III*. Harvard University Dept. of Linguistics.
- _____. 1991a. *Russian Sources on Korean Dialects*. Ph.D. dissertation, Harvard University.
- _____. 1991b. Korean Language Studies in the USSR: Past, Present and Future. *Icwungen.ehakhoy-ci* 8 (Journal of the Korean Society of Bilingualism), pp. 42-153. Seoul.
- _____. 1991c. A Soviet Korean Grammar from 1930. *Hankwukmal Kyoyuk* 3, pp. 153-178. Seoul.
- _____. (forthcoming-a) *The Korean Language in Imperial Russia*. Vol. 1: *Amateur Sources*. Vol. 2: *The Kazan' Materials*.
- _____. (forthcoming-b) *The Korean Language in the Soviet Union*.
- _____. (forthcoming-c). Archaisms and Innovations in Soviet Korean Dialects.

- Language Research*. Seoul.
- KPC. See Kim Pyengcey.
- KTk. See Kim Thay.kyun.
- Lee, Ki-moon [I Ki.mun]. 1961/1972/1983. *Kwuk.esa Kaysel* [Outline History of the Korean Language]. Seoul: Thap Chwulphan-sa.
- LCT. See 'Yu Changton.
- MLC. See Martin, et.al. 1968.
- Martin, S. E. 1993. *A Reference Grammar of Korean*. Tokyo: Tuttle.
- Martin, S.E., Yang Ha Lee and Sung-un Chang. 1968. *New Korean-English Dictionary*. New Haven: Yale University Press.
- Matveev, N.P. 1900. *Spravočnaja kniga g. Vladivostoka* [Reference Book for the City of Vladivostok]. Vladivostok.
- Pravoslavnoe missionskoe obščestvo [Orthodox Missionary Society]. 1904. *Opyt kratkago rusko-korejskago slovarja* [Attempt at a Short Russian-Korean Dictionary]. Kazan'. [Abbreviated as RKS].
- Putsillo, M.P. 1874. *Opyt rusko-korejskago slovarja* [Attempt at a Russian-Korean Dictionary]. St. Petersburg.
- Ramsey, S. Robert. 1978a. *Accent and Morphology in Korean Dialects*. Seoul: Thap Chwulphan-sa.
- Razvedčiku v Koree. (Rusko-korejskij slovar')* [For the Intelligence Officer in Korea. Russian-Korean Dictionary]. St.P.: Izd. Voenno-statističeskago Otdela Glavnogo štaba, 1904.
- RKS. See Pravoslavnoe...
- Se Cengmok. 1987. *Kwuk.e Uymun-mun Yenkwu* [Research on Interrogative Sentences in Korean]. Seoul: Thap Chwulphan-sa.
- _____. 1988. "Hankwuk.e Chengca Taywu Tungkup uy Hyengthaylon-cek Haysek (I) [A Morphological Analysis of the Politeness Levels in Korean]". In: *Kwuk.ehak* 17, pp. 97-152.
- Starčevskij, A.V. 1890. *Naši sosed. Spravočnaja knižka. Pograničnyj perevodčik...* [Our Neighbours. A Guide-book. A Border Interpreter...]. St. Petersburg.
- Xlynovskij, M. 1904. *Rusko-japonsko-korejskij voennyj perevodčik. 2-oe izdanie* [Russian-Japanese-Korean Military Interpreter. Second Edition]. Irkutsk.
- 'Yu, Changton. 1979. *Icoe Sacen* [Dictionary of Yi Dynasty Korean]. Seoul: Yensey Tayhakkyo Chwulphan-pu.
- Zhào, Xí and Xuān, Dēwǔ. 1986. "Cháoxiānyǔ Liùzhènghuà de Fāngyán Thèdiǎn [Dialect Characteristics of Korean 'Yukcin Speech] Mínzú Yǔwén 5, 1-13.

THE STATUS OF THE VOLGA-FINNIC LANGUAGES IN THE SOVIET COMPREHENSIVE SCHOOL

Seppo Lallukka
Neuvostoliittoinstituutti, Helsinki

1. Introduction. In Soviet educational history the initial period following 1917 can be described as one of radical experimentation. According to the Bolshevik schemes, the schooling of the rising generation was to be carried out through a unified labor school resting on the ideas of polytechnic education and on ambitious pedological innovations. At the beginning, the search for new forms allowed local variations in such matters as, for instance, curricula and teaching methods.

One additional feature of the new approach was the broad introduction of the minority languages in the educational system. However, in the course of the 1930s, at the latest, experimentation ended with a return to more sober practices. The language policy pursued in Soviet schools, as well, became subject to change. In spite of this, however, the use of local languages represents an element of the period of early educational reforms which has proved to be one of the most persistent.

The Finno-Ugrians of Russia were among those peoples who reaped benefits from the liberal language policy of the early Soviet regime. During this period of *korenizacija* (nativization), which lasted until the early 1930s, their languages were eagerly promoted. As an example, language standardization was notably intensified from what it had been earlier and the languages were increasingly used in publishing, in education, and even in administration. All this also applies to the Volga-Finnic branch of the Finno-Ugrian group, i.e., to the Mari and Mordvin languages.

The Mari and Mordvin homelands are located in the Middle Volga region. Since the Russian conquest of the Volga Valley, the territorial bases of these nationalities have been seriously eroded because of their own outmigration and also through massive immigration of other peoples, mainly Russians.¹ As a result the national areas of settlement

¹Under Soviet rule, autonomous territorial units were created for several Finno-Ugrian nationalities. Among these was the Mari Autonomous Oblast which was founded in November 1920 and elevated in December 1936 to the rank of an autonomous republic. It was obviously due to the highly dispersed settlement of the Mordvins that the buildup of Mordvin autonomy began somewhat later: The Mordvin *Okrug* was established in June 1928; at the beginning of 1930 the *Okrug* was transformed into an autonomous *oblast'* and later, in December 1934, into an

lost their consistency. One indicator of territorial dispersal is the fact that most Maris and Mordvins live outside their titular republics. The 1989 data show that not more than 27% of all Mordvins lived in the Mordvin Republic, and for the Maris the respective figure was 48% (NSS 1989:22-23; VS 1990:69).

In censuses taken after World War II, the Mordvins have shown continuing declines. Like the Mordvins, the Maris have also suffered assimilation losses but they, nevertheless, have been able to increase their numbers—if only modestly (Lallukka 1990:86, 247-52).

As can be seen from Table 1, there were 671,000 Maris and 1,154,000 Mordvins in the Soviet Union in 1989. Due to their linguistic assimilation (in effect, Russification), the number of native speakers of the Mari and Mordvin languages—542,000 and 774,000 respectively—falls considerably short of each nationality's total nominal strength. Furthermore, as regards the changes in the number of native speakers of the languages over the thirty years between the 1959 and 1989 censuses, speakers of the Mordvin languages (Erzia and Moksha) declined by 230,000 whereas the Mari linguistic community grew by 62,000. The period after 1970 shows, however, that the growth in the number of Mari speakers has come to a standstill.

The 1989 native-language retention rate calculated on the basis of Table 1 for all Mordvins is 67% and for the Maris, respectively, 81 %. What is of interest, though, is that within their titular republics the two nationalities show similar levels of native-language retention: for both of the nationalities the rate is 88 %.

However, one has to take into account that the census data on native language do not tell us to what extent the languages are used in practice or how fluent is their speakers' command. Since it is obvious that the census item on native language is interpreted by many people as a second question of ethnic affiliation, it is likely that, in the case of minorities like the Maris and the Mordvins, the census figures tend to underestimate the actual level of linguistic assimilation.

In this paper an attempt will be made to track the developments in the role accorded to the Volga-Finnic languages in Soviet comprehensive autonomous republic. After these two republics' declarations of sovereignty in 1990 the epithet "autonomous" was removed from their names.

TABLE 1

The Maris and the Mordvins: Number of population, number declaring the language of their own nationality as their first language and average annual growth rates (AAGR) during the intercensal periods; titular republic, outside the republic and total USSR, 1959-89:

| | Number (in thousands) | | | | AAGR (in percent) | | |
|---------------------------|-----------------------|-------|-------|-------|-------------------|-------|-------|
| | 1959 | 1970 | 1979 | 1989 | 59-70 | 70-79 | 79-89 |
| A. Population | | | | | | | |
| Maris | | | | | | | |
| Mari Republic | 279 | 299 | 307 | 324 | 0.6 | 0.3 | 0.6 |
| Outside Rep. | 225 | 299 | 315 | 347 | 2.6 | 0.6 | 0.9 |
| Total USSR | 504 | 599 | 622 | 671 | 1.6 | 0.4 | 0.8 |
| Mordvins | | | | | | | |
| Mordvin Rep. | 358 | 365 | 339 | 313 | 0.2 | -0.8 | -0.8 |
| Outside Rep. | 927 | 898 | 853 | 841 | -0.3 | -0.6 | -0.1 |
| Total USSR | 1,285 | 1,263 | 1,192 | 1,154 | -0.2 | -0.6 | -0.3 |
| B. Native speakers | | | | | | | |
| Maris | | | | | | | |
| Mari Republic | 273 | 287 | 287 | 287 | 0.4 | 0.0 | -0.0 |
| Outside Rep. | 206 | 259 | 252 | 255 | 2.1 | -0.3 | 0.1 |
| Total USSR | 480 | 546 | 539 | 542 | 1.2 | -0.1 | 0.1 |
| Mordvins | | | | | | | |
| Mordvin Rep. | 348 | 351 | 320 | 277 | 0.1 | -1.0 | -1.4 |
| Outside Rep. | 656 | 632 | 545 | 496 | -0.3 | -1.6 | -0.9 |
| Total USSR | 1,004 | 983 | 865 | 774 | -0.2 | -1.4 | -1.1 |

AAGR = $(1/k)\ln(P_i/P_0)100$, where k = number of years in period, P_0 = initial population, P_i = final population.

Sources: Itogi Vsesojuznoj perepisi naselenija 1959 g. RSFSR. Moscow: Gosstatizdat, 1963, table 54; Itogi Vsesojuznoj perepisi naselenija 1970 g. Vol. 4. Moscow: Statistika, 1973, tables 1, 3-4, 6; Čislennost' i sostav naselenija SSSR. Moscow: Finansy i statistika, 1984, tables 13-15; Nacional'nyj sostav naselenija. Vol. 2. Moscow: Goskomstat SSSR, 1989, pp. 22-23; Vestnik statistiki, 1990:10, p. 69.

school (*obščebrazovatel' naja škola*). The focus, thus, is on so-called national schools. These may be defined as schools whose students basically represent children of non-Russian nationalities and where their native language, to a varying degree, is used in the study process. The existing differences between Russian and national schools pertain to the peculiarities in the curriculum plan (Volkov 1989:33).

For a general survey of the status of the Volga-Finnic languages in schooling a good deal of quantitative information would be necessary on the scope of the national school network, the number of children attending such schools, in addition to matters of curriculum, i.e., the time allotted for study in the national language and/or for the study of the language as a separate subject.

The release of information on such issues had virtually stopped by World War II and it was not until the late 1980s that publication gradually resumed. The scattered notes which in the meantime could occasionally be glimpsed in Soviet publications were often contradictory and misleading. The gaps and defects in the data thus make it difficult to reach a balanced overview on the development of Mari and Mordvin national schools during the Soviet era. It turned out that there was somewhat more information available on the Mari than on Mordvin schools. This is also reflected in the following examination.

2. The Rise and Fall of the National Schools.

2.1. The Blossoming of *korenizacija*. During the second half of the 19th century the question of providing instruction in non-Russian languages began to surface from time to time in public discussions in the Russian Empire. Some practical steps were also taken toward that end, in particular, by the Kazan'-based missionary movement of the Russian Church (Kreindler 1979:5-26; Lallukka 1987:143-65). After the 1905 Revolution the debate became more intense, but by that time the whole issue had proved to be highly controversial. For the most part the initiatives to make teaching available in the national languages were turned down because in the ruling circles hostile attitudes prevailed toward any linguistic diversification of public life in the Empire (Pennar et al. 1971: 23-25, 163-64.).

Among the minorities of the Volga-Urals region, only the Tatars were able to establish schools of their own in the pre-Soviet period,

with subjects taught in Tatar even in the higher grades. These schools were financed by their own religious communities (Pennar et al. 1971:25; Volkov 1989:37).

The existing Mari and Mordvin schools were basically run in Russian but in some schools, especially in missionary ones, the native language was also used as an aid in teaching. For instance, as early as the 1870s instruction was provided in Mari at the initial phase in a number of schools (Volkov 1989:37-38; Stepanov 1989:55; Apakaev 1990: 7-8).

Acts favoring local languages were enacted by the Bolsheviks at the very beginning of their rule (Konstantinov 1958:6; Pennar et al. 1971:166-68, 192). The Narkompros (the RSFSR People's Commissariat of Education) stood at the head of the practical implementation of the new language policy in the school. Among the chief goals of the Bolsheviks was achieving universal literacy in the shortest possible time. It was also recognized that, regarding the minorities, the problem could be solved only by means of national schools and by providing instruction in the native tongue (Jaškin 1978:120). Consonant with this approach, the Narkompros issued a special decree in October 1918 "On the Schools of National Minorities," which stipulated that all nationalities were granted the right to organize instruction in their own languages, and that all minority schools were state schools (NO 1974:145).

A special section for the education of minorities functioned within the Narkompros with subcommittees for individual nationalities, each in charge of the guidance of education at the local level. For the Maris such a subcommittee was set up in 1919 and for the Mordvins the following year (Jaškin 1978:121; Kul'turnoe 1983: 23). About that same time the preconditions for the national school in the Mari and Mordvin areas were also being formed. These included, among other things, crash courses for teacher training and conferences for educational workers. Hand in hand with language standardization pursuits began work on new textbooks (Konstantinov 1958:6; Taldin 1958:139-44; Jaškin 1978 120-21; Kul'turnoe 1983:178-79).

There is obviously room for argument as to whether the early Bolshevik nationality policy was sincere or tactical, but for such nationalities like the Maris and Mordvins who were about to take their first steps toward modern nationhood the policy certainly opened unprecedented opportunities for cultural development.

On the other hand, the economic and cultural backwardness of the Volga-Finnic areas posed an obstacle to development. By Russian standards the Mari and Mordvin homelands lagged a great deal behind with respect to school facilities and the spread of literacy. Among the indigenous nationalities, in particular, the educational level was low (Pennar et al. 1971:17; Lallukka 1990:67). Moreover, the schools suffered from the Civil War and the famine of 1921-22 caused a further setback. Thus, it seems that in the principal Mari areas the school network was thinner in 1923 than it had been on the eve of Soviet rule (Kul'turnoe 1983:64; Ist. Mar. 1987:104).

The period from September 1922 to the end of 1925 marked a significant recovery in the Mari AO: over one hundred schools were reopened and more than 10,000 pupils resumed their studies (Ist. Mar. 1987:104). By 1926-27 the schools obviously had reached the level at which the coverage of children of school age had been at the beginning of the century (Apakaev 1990:11; Lallukka 1990:68). In the late 1920s and the early 1930s the Mari national schools advanced rapidly: the number of Mari primary schools in the AO rose from 159 in 1928-29 to 331 in 1932-33 (Očerki 1960:166-67, 211-12).

Before 1917, 89 schools intended for the Mordvins had been functioning in the territory of today's Mordvin Republic. By the middle of the 1920s the figure rose to about 200 and by 1929 to 335. The Mordvin languages were not used as the medium of instruction in all of them, though; around 1925, in the Krasnoslobodsk uезд, for instance, in 32 out of the total of 46 Mordvin schools teaching was carried out in the indigenous language (Taldin 1958:142-44; Lallukka 1990:68).

In spite of the considerable progress made in building up the public education system, a large portion of children of primary school age did not attend schools at the close of the 1920s. In the Mari AO the share of such children was 31% (for the Maris and Russians 37% and 25%, respectively) and in the Mordvin Okrug about half of the children—both Mordvin and Russian—did not go to school. Moreover, according to a Narkompros report of 1929, the bulk of national schools in the Middle Volga region functioned in poor buildings, mostly in peasant huts, and suffered from an acute shortage of textbooks (Pajmakov 1958:122; Jaškin 1978:126; Ist. sov. 1987:240-41). Thus, the Party decision of 1930 to introduce universal and compulsory primary education (the four-year school) made it necessary to further intensify

educational efforts. According to Soviet reports, by the middle of the 1930s practically all children of primary school age went to school in the Mari and Mordvin AOs (Jaškin 1978:125, 128; Očerki 1960:166-67, 211-12). To some extent the progress seems have been fictitious, however. Namely, it is evident from several sources that at the close of the 1930s and the beginning of the 1940s there were still gaps in the coverage of primary school (NO 1974:116-17; Akapaev 1990:25; Kul'turnoe 1983:172).

Another decision made in 1930, which was directly related to the national school, was made by the council of the Narkompros. It stipulated that by the 1931-32 school year each primary school should change over to the native language of the majority of its pupils. This part of the decision obviously was realized for the most part, but analogous provisions concerning language *korenizacija* in the secondary school by 1932-34 remained very far from being carried out in practice (Simon 1986:68).

The network not only of primary schools but also of secondary schools grew intensively, especially in the 1930s, after the introduction of compulsory primary education (Taldin 1958:146). Along with this expansion the Volga-Finnic languages, as well, were able to secure a certain foothold in the secondary schools. Thus, the use of the Mari and Mordvin languages grew notably in grades 5 through 7, and during the second half of the 1930s, schools also existed which were operating in these languages up to the 10th grade (BSE 1938:130-31; Pennar et al. 1971:196; Kreindler 1984:16).

What percentage, then, of the Volga-Finnic children were taught in their native languages? The data for 1927 indicate that 43.5% of the Mari children attending primary schools of the RSFSR were taught in Mari. For the Mordvins the respective figure was 19.4%. A large share of children were also attending primary schools with more than one language of instruction. In regard to the Volga-Finnic nationalities this meant, practically with no exception, that it was Russian, specifically, that was used along with the native language. Such mixed schools were attended by 43.6% of the Mari and 42.6% of the Mordvin pupils. Instruction was entirely in a non-native language (overwhelmingly in Russian) for 12.5% of the Mari and 37.3% of the Mordvin children enrolled in primary school (Nacional'naja 1930:278).

It would thus seem, on the basis of the 1927 figures, that the Mari

children were better provided with native language schools than the Mordvin children. One has to take into account, however, that the figures refer not just to the ethnic homelands of the two nationalities but to the entire RSFSR. Since the Mordvins live more dispersedly than the Maris—and because national schools are usually to a lesser extent available for diaspora groups—it is not surprising that according to the RSFSR-wide figures the Maris compare favorably with the Mordvins.

Another set of data refer to the Mari and Mordvin ASSRs at the beginning of the 1938-39 school year. It turns out that in the Mari Republic 29% of all school children were taught in Mari and in the Mordvin Republic the analogous figure was 34.3%. The Mordvin figure corresponds rather well to the share of Mordvin population in the republic (37.4% in 1939), but for the Maris there is a discrepancy of 18 percentage points, since the titular nationality made up 47.2% of the population in the Mari ASSR in 1939 (Pennar et al. 1971:196; Kozlov 1960: 62; BSE 1954:276).

The same data describe the situation also by grades. There was practically no change in the percentage of children obtaining instruction in the Mordvin languages when moving from grades 1-4 to grades 5-7 and then to 8-10. In all cases the figure was 34 to 35% of all students attending the respective group of grades. In the Mari ASSR the picture was dissimilar since the figures declined clearly toward the senior grades: in grades 1-4, 34.5% of children were taught in Mari; in grades 5-7 and 8-10, the figures were respectively 19.2% and 5.4%. The discrepancy with the overall composition of population thus grew quite striking in secondary school (Pennar et al. 1971:196).

In their titular republic, the Mordvins were obviously quite successful in rooting instruction in native language through the entire comprehensive school. This is quite a notable achievement because in most minority areas of the RSFSR the secondary school, particularly grades 8 through 10, remained heavily dominated by Russian in the 1930s (Simon 1986:70-71).

2.2. The Fading of *korenizacija*. Elements of backsliding in the Bolshevik nationality policy toward more centralist practices began to emerge in the late 1920s. Some years later it was declared openly that local nationalism—instead of (Russian) great-power chauvinism—represented the main danger in the field of national relations.

In regard to language policy, a mixed picture emerges of the developments in the Middle Volga region. On the one hand, the *korenizacija* of the local languages continued to make progress in the 1930s in several fields, perhaps most notably in the secondary school. On the other, a growing emphasis on Russian was also felt. To some degree this was obviously motivated by economic grounds. The development of heavy industry as a unified USSR-wide complex provided a rationale to switch to teaching in Russian in many higher educational institutions. One can assume that somehow this must have had repercussions also on the comprehensive school.

Formally, the repressions of the 1930s did not touch the Volga-Finnic national school, but in practice severe damage was inflicted. First, not only politicians and functionaries of high culture but also a great number of teachers were silenced or liquidated as "bourgeois nationalists." Among them, those who taught native language and literature were particular targets. Second, a great number of textbooks, especially those not printed in Russian, were destroyed (Kreindler 1984:18-20; Sanukov 1990:86).

In the Mari Republic the repressions were so penetrating that in a number of schools it was not possible to get the school year started in the autumn of 1937 because of the lack of teachers (Sanukov 1990:86). A Mari historian portrays the situation with the following gloomy lines:

The disaster fell, above all, on the teachers of the Mari language and literature. And if someone was left untouched, it was only possible to use newspapers for the lessons because all textbooks and all books in Mari had been confiscated.

...

In February 1938, with growing anxiety, the department of public education complained to the Narkompros of the RSFSR about the situation: as a result of acts of sabotage, schools in the republic were left without programs, textbooks, and auxiliary material on the Mari language and literature; those which had existed earlier had been produced by counter-revolutionaries and there was nobody who could cope with the task of writing new ones. (Sanukov 1990:86-89)

It is more than obvious that when the production of textbooks later, nevertheless, did resume, the new books were deprived of almost any national content.

In March 1938 the Party and the Soviet government issued a decree which made the Russian language a compulsory subject in all non-Russian schools. The decree certainly reflected the growing importance laid on Russian,² but as such it did not necessarily impede the use of local languages in national schools. The decree of 1938 even made the idea quite explicit that a child must be taught in his or her native tongue:

the native language forms the basic medium of instruction in the schools of national republics and oblasts; the deviations from this rule that take place in some autonomous republics of the RSFSR are permissible only as provisional arrangements; the tendency of changing the Russian language from a subject of study to the language of instruction, which, in fact, suppresses the native language, is harmful and incorrect. (Sovetkin 1958:15)

Strictly speaking, however, the introduction of Russian as a compulsory subject was a departure from the principle of language equality. Russian was acknowledged as being more important than other languages. In view of its functions, however, one can say that Russian had always been of more importance (Kirkwood 1991:63-64). On the local level the message conveyed by the decree of 1938 often tended to be interpreted as a directive toward emphasizing Russian at the cost of the indigenous language. The immediate shifts, however, were not necessarily large and the evidence of changes in the relationship between Russian and the local languages is often conflicting. One can find evidence of continuing efforts to advance language *korenizacija* also in the period following 1938.

In February 1941, for instance, the Mari obkom of the Party passed a decree which, along with an appeal for improving the quality of Russian instruction, very explicitly set the goal that all children in Mari primary classes be taught in Mari in the 1941-42 school year. Furthermore, it was also noted that all those who leave Mari primary schools should be taught in the native language while studying in the senior grades (Kul'turnoe 1983:173).

²The requirements set for non-Russian pupils for a knowledge of Russian were quite high: "Students leaving incomplete secondary school (seven-year school) must be able to express their thoughts in Russian freely and correctly, both in speech and in writing" (Krasnov 1972:80).

However, a gradual shift toward Russian was inevitable. This trend was reinforced by increased population mobility connected with the war, postwar reconstruction, urbanization, and migration. Intensified administrative centralization also added to the importance of Russian, since without a sound knowledge of that language advancement in one's career became increasingly difficult.

Most Mordvin schools outside the titular republic were closed in the 1940s and the 1950s (Kreindler 1984:18). The same pattern—although not as far-reaching a one—obviously fits the Mari schools outside the Mari ASSR. One can also conclude that during the same decades the Mari and Mordvin languages somehow lost their position as languages of instruction in grades 8 through 10. It turns out namely, that on the eve of Khrushchev's school reform, they were not being used as mediums of instruction other than during the first seven years of schooling (Sovetkin 1958:2).

2.3. The Decline Turns into a State of Decay. The Education Laws of 1958-59, which launched Khrushchev's school reform, opened the way for sharp reductions in the provision of native-language schooling for most nationalities of autonomous republic standing or of lesser political status.³ At first glance, it would appear that the reform restored the equality of languages by seemingly abolishing the obligatory study of Russian and by granting parents the right to select the language in which their children would be instructed. In actual practice, however, the reform turned out to be "a form of democratic legitimation for reductions in the provision of native-language schooling through the exercise of 'free choice' by the parents of non-Russian children (Anderson et al. 1984:1021)." With typical Soviet doublespeak, Article 15 of the RSFSR Educational Law of 1959 declares:

It is decreed that instruction in the schools is to be conducted in the native language of the pupils. The parents are granted the right to decide to which school, with which language of instruction, to send their children.

In schools in which instruction is conducted in the language of an autonomous republic, autonomous oblast, or national okrug, the study of Russian is to be conducted according to the wishes of the pupils;

³For overviews of the language issue in the educational reform, see Bilinsky 1962:138-57 and Simon 1986:280-87.

likewise in Russian schools the pupils, if they so wish, can study the language of the autonomous republic, autonomous oblast or national okrug. (Vedomosti 1959:405-06.)

Later on, the matter was made more explicit in all-Union documents. Article 20 of the USSR Law on Public Education enacted in 1973 states that the pupils of general education schools are provided with the possibility "to study in their native language *or in the language of another people of the USSR*" (emphasis added) (NO 1974:97). Quite clearly, the idea that a child must be taught in the native language, which was clearly stipulated in the decree of 1938, was pushed aside.

Given the dominant role of the Russian language in the society, the voluntary choice of the language of instruction has hampered the learning of the indigenous languages. In the RSFSR this stems from the unavailability of educational facilities in non-Russian languages beyond the national school. Therefore, many parents tended to reason that by opting for a non-Russian school for their children they would actually place their children at an unnecessary disadvantage vis-à-vis the pupils of Russian schools with regard to their prospects for entering college and other higher educational institutions. Since the native language appeared to lack a future, many non-Russian parents were obviously quick to ask for exemption for their children from studying the native language even as a subject.

On the surface, certainly, the right granted to the parents to select the language of instruction for their children would seem to present a democratic innovation. But even if one considers the measure an extension of democratic rights, the problem still remains that the practical implementation of this section of the school reform often proceeded with little attention to parental wishes (Lallukka 1990:185). Thus, in a situation where school graduates have an opportunity to continue their studies only in Russian, everything worked against the minority languages. The way in which Russian and the local languages were placed against one another must be considered an outright Russificatory measure that seriously undermined the functional scope of languages such as the Volga-Finnic.

In pedagogical terms the switch from the native language of instruction to Russian was repressive. In rural areas the pupils' command of Russian was very often not good enough for pursuing studies in that

language. Moreover, there was a shortage of teachers adequately prepared to make the switch in the language of instruction easier for the children. The general level in the quality of knowledge suffered (Kirdjaškin 1973:77-80; Gruzov 1989:81-82; Kirkwood 1991: 64). Up to the present the so-called direct method (*prjamoj metod*), i.e., the practice of pursuing studies completely in Russian from the very beginning, continues to pose a pedagogical problem in the schooling of minority children (Sukunov 1989:60-63; Rekomendacii 1989:116).

How, exactly, the reductions in the provision of native-language instruction were administered can be traced only approximately. At any rate, the point of departure is that the Mari and Mordvin languages served as mediums of instruction in the first seven grades. In June 1959 the Supreme Soviet of the Mari ASSR enacted a decree stipulating that beginning from the fifth grade, national schools were to function in Russian. The implementation of the switch began from the 1959-60 school year. At the same time, an analogous step was taken in the Mordvin ASSR, as well.

Several republican leaders also made it quite explicit where their sympathies lay on the language issue. For instance, the chairman of the Council of Ministers of the Mari Republic, G. I. Kondrat'ev, spoke in public: "I am against the teaching of the Mari language in schools" (Veden'kin 1990:83).

There are some inconsistencies among the sources as to whether the ouster of the native language from the Mordvin schools immediately included the fourth grade also (Kotok 1964:89; Gusev 1971:50; Kirdjaškin 1973:77). But by 1972, at the latest, the cutbacks had gone so far that no more than the first three grades of primary school were available in Mari and Mordvin. Beginning from the fourth grade all instruction was in Russian (Danilov 1972:23). As a separate subject, though, it was possible to continue the study of the native tongue in the senior grades as well.

In the 1960s and the 1970s there was also a drive to rationalize the rural school network, which meant bigger schools and drawing children from a larger geographical area. The schools became ethnically more heterogenous, which advanced the use of Russian as a lingua franca.

A change which has occurred since the 1970s is that in some schools instruction is now provided in Mari or Mordvin up to and including the fourth grade. This is connected with the transfer from a 10-year

school to an 11-year school, with children starting school at the age of six, in accordance with the school reform of the 1980s (Gruzov 1989:82).

3. The Current Situation (the 1980s)

3.1. The Curriculum. During Khrushchev's and Brezhnev's periods of reign, the Volga-Finnic languages thus dropped to second rate status in the comprehensive school. The enrollment in the national schools, a topic to be discussed later, also underwent a progressive decline. All this is in line with the aggressive language policy of those years: the ever-increasing importance of Russian was continually pushed forward and it was trumpeted as the "second native language" for all non-Russians. The role of Russian also went far beyond the functions of a lingua franca: it increasingly became an ideological force, a fundamental carrier of the Communist message and an agent of a "new historical entity, the Soviet People" (Kirkwood 1991:65).

Not surprisingly, the boosting of Russian was to be felt in the national school as well. The questions of intensifying the study of Russian were widely discussed at frequent all-Union and local conferences devoted to Russian language in schooling.

At the present time two basic types of national schools exist in the Volga-Finnic homelands:

(1.) Schools functioning on the basis of the so-called combined method (*kombinirovannyj metod*), i.e., the indigenous language is the main medium of instruction in the first three or four grades. Beginning with the fourth or fifth grade, instruction is carried out in Russian and the indigenous language is taught as a subject.

(2.) Schools applying the so-called direct method (*prjamoj metod*) which means that Russian is used as the medium of instruction from the very beginning. According to parental wishes, the indigenous language is studied in these schools as a separate subject.

Intermediate types also exist. For instance, in type (1.), the switch to Russian can also occur earlier than after three years (Kotok 1964:89; Gruzov 1989:82). Quite a widespread practice is to gather the children

of different nationalities under the same roof, in schools in which some groups are taught in Russian and others—at the beginning—in a Volga-Finnic language.

An overview of the number of hours devoted weekly to the study of the Russian language, literature, and native language throughout the entire school program can be gained from the Appendix. In both types of schools the obligatory study of Russian starts in the first grade. The study of the native tongue in type (2.) schools begins in the first grade and the number of hours allotted for this subject varies from two to three hours per week, depending on the grade. One must note, however, that according to press reports, the indigenous language in type (2.) schools represents, in practice, a voluntary subject.⁴ In type (1.) schools the weekly number of hours mandated to the native language is six or eight during the first school year, then the amount gradually declines approaching the senior grades.

In all type (2.) schools and in the senior grades of type (1.) schools, most subjects are taught using Russian textbooks prepared for all-RSFSR use. The main exceptions from this rule are the native language and literature, as well as Russian language and literature. Complaints have been made, however, that in many national schools of type (2.) Russian language is being studied according to the programs and textbooks prepared for Russian schools. This obviously causes additional problems for non-Russian pupils. It also turns out that some national schools, including Volga-Finnic ones, are national in name only, since in some schools the native language is not studied at all for lack of teachers, negligence on the part of the educational officials, or some other reason (Kirdjaškin 1973:78-80; Sabatkoev 1989:20; Sukunov 1989:62-63).

3.2. The National Schools: Network and Attendance. The discussion of the postwar period has so far concerned just the existence of certain types of national schools with no reference to the question of how readily the various educational facilities are within the reach of the Volga-Finnic children and—given the principle of free choice—to what extent the parents have enrolled their children in such schools.

⁴Janalov, V., "Na jazyke ravnopravija." *Molodoj kommunist*, May 24, 1989; Terexina, M., "Čto imeem—ne xranim, poterjavši—plačem," an article from a newspaper of the Mordvin Republic in 1990, the name of the paper could not be established.

Besides the types of national schools discussed above there are Russian schools in the Volga-Finnic homelands functioning entirely like schools in the Russian areas of the RSFSR and offering no possibilities for the study of the indigenous language of the area. As a matter of fact, these schools form the largest group in terms of the number of students enrolled (Narodnoe 1989:163; Mokšin 1989:131). A large segment of Mari and Mordvin children attend Russian schools. In principle, it is also possible that Russian children go to Volga-Finnic national schools but in actual practice the crossing of ethnic boundaries in this direction seems to happen rarely. Some Tatar schools also exist in the Mari and Mordvin republics (Mokšin 1989:131; Programma 1990:37, 39) and it is likely that some non-Tatars attend them. Thus, among the students of the Volga-Finnic national schools there obviously are some non-Maris and non-Mordvins but, given the asymmetry in the pattern of how ethnolinguistic circles are crossed, their proportion must remain slight.

In Soviet sources the various types of national schools are often lumped together and, as was noted earlier, many schools are only nominally national. Moreover, the mere number of schools does not help much to gain a picture of the proportion of children who are taught in the indigenous language or who study it, because the schools vary greatly in size. When one adds to this the fact that the available information is meager and scattered, it becomes clear that it is possible to reach not more than an approximate view of the network and attendance of the national schools.

3.2.1. The Titular Republics. Like ethnic cultural support in general, the provision of schooling in Volga-Finnic languages is, to a large measure, concentrated in the respective republics. Only some of the most significant diaspora groups outside the titular republics have been granted some kind of national schools.

In the Mari and Mordvin republics themselves, too, the native-language school network has never been complete enough to provide every school-aged member of the indigenous nationality with a realistic possibility to opt for a national school program. Since the 1958-59 reform, the educational opportunities offered in the Volga-Finnic languages have, as a rule, narrowed continuously. Present-day urban Mari and Mordvin children overwhelmingly attend Russian schools. It even appears that, with a few exceptions, type (1.) schools exist, in reality, only in the countryside. Until quite recently, Ioshkar-Ola and Saransk

were republic capitals without even the possibility to study the language of the titular nationality as a subject at school.

In the 1989-90 school year there were about 430 schools in the Mari Republic, 248 of which were Mari national schools. The Mari language served as the medium of instruction in the primary grades of 191 schools; of these 145 were Meadow Mari schools and the rest, 46, Hill Mari schools. Type (2.) schools numbered 57 (Meadow and Hill Mari was taught in 49 and 8 schools, respectively).⁵

In the same school year, the total number of children taught in Mari in grades 1 to 3, or 1 to 4, of type (1.) schools was 7,800. In type (2.) schools and in the senior grades of type (1.) schools altogether 18,000 pupils studied Mari as a subject. Summing these two figures and knowing that Mari children attending comprehensive schools totalled about 52,000, it is possible to conclude that some 50% of all Mari pupils in the republic were exposed to the Mari language at school, either in the form of a subject of study or in the form of the language of instruction.⁶ In other words, about half of the Mari children were enrolled in Russian schools.

It was not possible to establish how the 52,000 Mari pupils were distributed by grades in comprehensive school. However, supposing that the breakdown by grade follows the general Soviet pattern,⁷ it would seem that some 35 to 45% of Mari children attending schools in the republic were obtaining their primary schooling in Mari at the close of the 1980s.

The data on the Mordvin Republic of the early 1970s put the number of Mordvin schools at 391; most of them, 235, were Moksha schools and the rest, 156, were Erzia schools. The schools were attended by 41,000 Moksha and 36,000 Erzia children. Altogether these 77,000 children represented 36.9% of all school children in the republic (Kirdjaškin 1973:72; Mokšin 1977:201). Obviously, this figure falls somewhat short of the Mordvin percentage among children of school age in the republic (Cf. Lallukka 1990:152). The difference, however, was not very large.

⁵ Unpublished data in author's possession.

⁶ Unpublished data in author's possession.

⁷ In 1989-90, 37% of pupils attending comprehensive school in the Soviet Union were in grades 1 through 4. Cf. Narodnoe 1990:194. Applying this percentage to the Mari data produces an estimate of some 19,000 children attending primary schools.

Examining the above figures, it is important to note that by the early 1970s, type (1.) had been overtaken by type (2.) schools. Less than 10% of Mordvin pupils went to schools with Erzia or Moksha as the medium of instruction in the primary grades (Mokšin 1977:201).

In 1989 there were 289 Mordvin national schools left in the Mordvin Republic (180 Moksha and 109 Erzia schools out of a total of 839 schools). The number of primary schools offering some instruction in Mordvin was apparently somewhat less than 60 (cf. Mokšin 1989:131; Narodnoe 1989:163; Mišanin 1990:39).⁸ During the 1980s the number of children learning Mordvin as a subject fell from 80,000 to 24,000 (Mišanin 1990:39.).

Indicative of the continuing loss of ground among the Mordvin languages is also the fact that the number of children taught in Erzia or Moksha plunged from the 1980-81 figure of 12,000 to 4,000 in 1988-89.⁹

On the basis of the above information one would venture to say tentatively that approximately one half of Mordvin children were in some way in touch with the Erzia or Moksha language at school. Seen superficially, this proportion seems to be rather close to the analogous figure which was calculated for the Maris of the Mari Republic. In essence, however, there is a big difference: type (1.) schools predominate among the Mari national schools whereas the Mordvin schools overwhelmingly represent type (2.). In this important respect, i.e., as a medium of instruction, the position of the Mari language is much stronger than the position held by the Mordvin languages. Moreover, it seems that the reductions which took place in the number of students attending national schools during the 1980s were felt much more heavily by the Mordvins than the Maris (cf. NO 1989:88).

3.2.2. The Mari and Mordvin Diaspora. The Eastern Maris constitute the largest Mari diaspora group with their number now apparently exceeding 160,000. They do not form a coherent entity, however, but consist of several areal clusters. The largest number of Eastern Maris, 106,000 people, live in the Bashkir Republic (Lallukka 1990:113,

⁸There is disagreement between sources to what degree the Mordvin languages generally serve as mediums of instruction. Some sources speak of the first two grades whereas some of grades 1 to 3. Cf. Volkov 1989:34; Šaljaev 1977:63.

⁹The figures refer to the entire RSFSR but in fact all Mordvin schools representing type (1.) seem to be located in the titular republic (NO 1989:88; Mokšin 1989:131).

120). There are also Mari schools in Eastern Mari areas.

In the whole RSFSR, children taught in the 1988-89 school year in Mari totalled around 12,000 (NO 1989:88). Subtracting from this the pupils taught in Mari in the Mari Republic (7,800) one gets an approximate figure of 4,000 for the number of Mari children enrolled in type (1.) schools and instructed in Mari outside the titular republic. Most of these children were likely to be residents of Bashkiristan, but in the Tatar and Udmurt republics, as well, there are some Mari schools of type (1.). In these republics some type (2.) Mari schools are also functioning (Programma 1990:37-40).

How many Mari schools there are in all in the Eastern Mari areas could not be established. The bulk of them obviously are in Bashkiristan, where there are about 150 of them. Besides this it is only known that there are 13 Mari schools in Sverdlovsk oblast, all of which represent type (2.).¹⁰

Closer to the Mari Republic, in Kirov oblast, is found a relatively sizeable Mari population of 44,000 people (Sanukov 1992:25), but until recently, at least, they have been educated only in Russian.¹¹ In consequence, these people are generally illiterate in Mari (Gruzov 1989:83). What the public opinion on this matter is remains unknown to an outsider, but at the present time, one can come across voices grieving over the fate of the Mari language in those areas:

I am a 14-year-old Mari girl, but the Mari language I know only in speech and even then I cannot find all the words. I live in the Kiknur district of the Kirov oblast where most of inhabitants are Maris but nevertheless, not a single national school exists here in the district. Many children no longer speak Mari because they are ashamed and I, as well, am ashamed of the native tongue. I think that some time will pass and the Mari language will not be known at all. Why is this so? I want to know my language! (From a letter read by A. Ivanova at the First International Congress of Finno-Ugrian Writers, Ioshkar-Ola, May 22-27, 1989.)

Much less can be said about Mordvin schools. According to some sources there were no Mordvin national schools, at all, outside the

¹⁰ Unpublished data in author's possession.

¹¹ In 1990 two schools in the Mari areas of Kirov oblast went over to some sort of national program. Obviously, they started to teach the Mari language as a subject in these schools. See Sanukov 1991:75.

titular republic in the late 1980s (Xuzangaj 1988:219-20; Mokšin 1989:131). However, there is some obscurity about this matter and on the basis of one document one can infer that there could be some Mordvin schools at least in Bashkiristan and Tataristan (Programma 1990:37-40).

What seems clear, however, is that, even if Mordvin schools exist outside the Mordvin Republic, they can only form an extremely thin network. One can also conclude that the non-republic Mordvins are in a much worse position than their Mari counterparts regarding their possibilities to be educated in national schools. It goes without saying, as well, that the coverage of both the Mari and Mordvin diasporas by national schools lags far behind what it is found, respectively, in each of the titular republics. Because of the large diaspora sections, the overall coverage of the Maris and Mordvins by national schools remains much more incomplete than what it would seem on the basis of first impressions founded on data pertaining only to the titular republics. This especially holds true of the Mordvins.

4. Conclusion. By and large, the role assigned to the Mari and Mordvin languages in the Soviet comprehensive school has been subject to those shifts which have taken place in the arena of the general language policy in the country. At the beginning of Soviet rule, the Volga-Finnic languages were actively promoted. Then their favoring gradually ended and, beginning with the late 1950s, there was a distinct decline which relegated the languages into a secondary role, whereas Russian was pushed eagerly forward at the same time. One can say that there is an intrinsic conflict in Soviet bilingual education because some look upon it as a means of promoting Russian, while others see it as a safeguard for minority languages.

Taken in their entirety, both Mari and Mordvin children were mostly educated in Russian in the late 1980s. In their titular republics, though, about half of the pupils of each Volga-Finnic nationality were in touch with the indigenous language at school: they either learned it as a subject or were taught in it in the primary grades. As a medium of instruction Mari has been able to keep its position considerably better than the Mordvin languages. This is, however, far from suggesting in any way that the status of the Mari language in school would be satisfactory from the point of view of language maintenance.

Certainly, in the absence of school facilities which would operate in the mother tongue, the right granted the parents to choose the language of instruction for their children is not much more than a nominal and fictional right. It happens in the Mari and Mordvin republics that in school districts with a population consisting of 80 to 90% of the indigenous nationality all children are taught in Russian right from the beginning (Gruzov 1989:83). Most recently, the principle of parents' choice was legally reaffirmed in the USSR Law on Languages of 1990 (Vedomosti 1990:19). Thus, in this regard there seems to be no principal change in language policy.

What has changed is the public discussion being carried on in the republics. Since the second half of the 1980s, the alienation of the Volga-Finnic peoples from their languages has become a frequent topic of public debate. Moreover, it appears that with the birth of movements for national awakening, the discussion has assumed a rather straightforward nature. It is increasingly being realized that a balanced view cannot be achieved without taking into account the realities of the context, and in particular, the part played by Russian in the decline of the Volga-Finnic languages.

Setting forth national-Russian bilingualism as an aim of language policy among non-Russians represents a key link in Soviet language policy. And there is no denying that under present-day conditions a working knowledge of Russian is a practical necessity for virtually all residents of the Middle Volga region. However, from the point of view of the maintenance of the Volga-Finnic languages, national-Russian bilingualism poses a serious problem, since it has proved not to be a stable state but basically a transitional stage on the road toward Russian monolingualism. It is, therefore, not surprising that in current debates one often comes across appeals for "equal," "harmonious," or "symmetrical" bilingualism (Rekomendacii 1989:117; Rezolucija 1990:11). The admission has also been made that type (2.) schools using the so-called direct method do not offer support for national languages which would allow them to compete fairly with Russian (Gruzov 1989:84; Volkov 1989:33-34). For strongly Russified pupils the study of the indigenous language at school has increasingly become similar to learning a foreign language.

On the other hand, type (1.) schools, as well, are insufficient in a situation where a formidable proportion of Volga-Finnic children have

a poor command of their indigenous language and are not able follow instruction in that language. Overall, there are no easy solutions which would make possible a reversal in the current declining trend of the languages.

The opportunities available during the last few years to speak out about the problems of the Volga-Finnic languages have obviously not been without their consequences. Namely, educational authorities have begun to seek ways for improvement. A recent program of the RSFSR Ministry of Education itemizes a number of measures planned for implementation during the first half of the 1990s in national schools. Among others, these include the following: in all Mari and Mordvin schools of type (1.) instruction in the indigenous language is going to be extended to include grades 1-4, there is a plan to increase the production of textbooks in the Mari and Mordvin republics by some 50 to 100% from the level of the preceding 5-year period; and an increase in the student intake is also expected. Most of this increase, though, is aimed for mixed (multinational) schools of the Mordvin Republic which offer the possibility to study Erzia/Moksha only as an optional subject (Programma 1990:26, 37, 44).

The national school constitutes one of the most important language maintenance institutions for the Volga-Finnic peoples. It is therefore not surprising that the progressive declines in the provision of native-language schooling have been followed by the rising linguistic Russification of the nationalities, in particular, of their younger generations. The countermeasures of the last few years will hardly be able to reverse the decline of the languages. One can hope that these measures are but harbingers of a changing approach to the whole matter, which in the future would allow a comprehensive recovery of the national school.

It is evident enough that the present-day Mari and Mordvin national schools fail to offer a basis for a stable bilingualism and language maintenance. All too often, however, in matters like this, all the blame is poured on the Russians. One should remember that influential posts in the local educational administration have often been held by ethnically homeless, marginal, minority descendants who have easily ignored the specific educational needs of the nationality of their origin. In conclusion it should also be noted that the Soviet treatment of minorities is too often measured against some idealized and abstract standard of fairness,

instead of comparing it with the performance in other states. It is true that the native-language education of the Volga-Finnic minorities is in poor standing, but the fact that the languages nevertheless continue to have a footing, even if narrow, in school tells of a readiness on the part of the Soviet government to perpetuate a certain ethnolinguistic diversity in the country. Not all governments have been willing to do that.

REFERENCES

- Anderson, B. and Silver, B. 1984. "Equality, Efficiency, and Politics in Soviet Bilingual Education Policy, 1934-1980." *American Political Science Review*, 78:4.
- Apakaev, P. A. 1990. *Prosvetiteli marijskogo kraja*. Joškar-Ola: Marijskoe kn. izd-vo.
- Bilinsky, J. 1962. "The Soviet Education Laws of 1958-9 and Soviet Nationality Policy." *Soviet Studies*, 14:2.
- BSE. 1938. = *Bol'shaja sovetskaja ènciklopedija*, 1st edition, vol. 36. Moscow: Sovetskaja ènciklopedija.
- BSE. 1954. = *Bol'shaja sovetskaja ènciklopedija*, 2nd edition, vol. 26. Moscow: Bol'shaja sovetskaja ènciklopedija.
- Danilov, A. 1972. "Mnogonacional'naja škola RSFSR— praktičeskoe voploščenie leninskoj nacional'noj politiki." *Narodnoe obrazovanie* 12.
- Gruzov, L. P. 1989. "Marijsko-russkoe dvujazyčie i voprosy soveršenstvovanija prepodavanija russkogo jazyka v škole." *Aktual'nye problemy literaturnogo jazyka*. Joškar-Ola: Marijskij NII.
- Gusev, D. I. 1971. "Razvitie kul'tury narodov Povolž'ja i Prikam'ja v period stroitel'stva kommunizma." *Voprosy istorii* 4.
- Ist. Mar. 1987. = *Istorija Marijskoj ASSR*. Vol. 2. Joškar-Ola: Marijskoe kn. izd-vo.
- Ist. sov. 1987. = *Istorija sovetskogo krest'janstva Mordovii*. Vol. 1. Saransk: Mordovskoe kn. izd-vo.
- Jaškin, I. A. 1978. *Mordovskaja socialističeskaja nacija—detišče Oktjabrja*. Saransk: Mordovskoe kn. izd-vo.
- Kirdjaškin, V. V. 1973. *Mordovija na puti k vseobščemu srednemu obrazovaniju*. Saransk: Mordovskoe kn. izd-vo.
- Kirkwood, M. 1991. "Glasnost', 'The National Question' and Soviet Language Policy." *Soviet Studies*, 43:1.
- Konstantinov, N. A. 1958. "Osnovnye voprosy stroitel'stva nacional'noj školy za 40 let Sovetskoi vlasti." *Nacional'nye školy RSFSR za 40 let*. Moscow: Izd-vo Akademii pedagogičeskix nauk.
- Kotok, E. 1964. "Mordovskie školy v novom učebnom godu." *Russkij jazyk v*

- nacional'noj škole 6.
- Kozlov, V. I. 1960. "Rasselenie mordvy." Trudy Instituta ètnografii AN SSSR, 63.
- Krasnov, V. 1972. "Jazyk mežnacional'nogo obščeniia." Narodnoe obrazovanie 12.
- Kreindler, I. 1979. "Nikolai Il'minskii and Language Planning in the Nineteenth-Century Russia." *International Journal of the Sociology of Language*, 2.
- _____. 1984. *The Mordvinians—A Doomed Soviet Nationality?* The Hebrew University of Jerusalem. The Soviet and East European Research Centre. Research Paper no. 56.
- Kul'turnoe. 1983. = Kul'turnoe stroitel'stvo v Marijskoj ASSR. 1917-1941 gg. Joškar-Ola: Marijskoe kn. izd-vo.
- Lallukka, S. 1987. "Kazan' Teacher's Seminary and the Awakening of the Finnic Peoples of the Volga-Urals Region." *Studia Slavica Finlandensia*, 4.
- _____. 1990. *The East Finnic Minorities in the Soviet Union. An Appraisal of the Erosive Trends.* *Annales Academiae Scientiarum Fennicae*, Ser. B, Vol. 252.
- Mišanin, Ju. A. 1990. "Nacional'naja pečat' i nacional'naja kultura: vzaimodejstvie i vzaimovlijanie." *Vestnik Mordovskogo universiteta* 3.
- Mokšin, N. F. 1977. *Ètničeskaja istorija mordvy.* Saransk: Mordovskoe kn. izd-vo, 1977.
- Nacional'naja. 1930. = Nacional'naja politika VKP(b) v cifrax. Moscow: Izd-vo Komakademii.
- _____. 1989. *Mordovskii ètnos.* Saransk: Mordovskoe kn. izd-vo.
- Narodnoe. 1989. = Narodnoe xozjajstvo RSFSR v 1988 g. Moscow: Finansy i statistika.
- Narodnoe. 1990. = Narodnoe xozjajstvo RSFSR v 1989 g. Moscow: Finansy i statistika.
- NO. 1974. = Narodnoe obrazovanie v SSSR. Sbornik dokumentov 1917-1973 gg. Moscow: Pedagogika.
- NO. 1989. = Narodnoe obrazovanie i kul'tura v SSSR. Moscow: Finansy i statistika.
- NSS. 1989. = Nacional'nyj sostav naselenija. Vol. 2. Moscow: Goskomstat SSSR.
- Očerki. 1960. = Očerki istorii Marijskoj ASSR (1917-1960 gg.). Joškar-Ola: Marijskoe kn. izd-vo.
- Pajmakov, A. S. 1958. "Razvitie škol'nogo obrazovanija v Marijskoj ASSR." Nacional'nye školy RSFSR za 40 let. Moscow: Izd-vo Akademii pedagogičeskix nauk.
- Pennar, J., Bakalo, I., and Bereday, G. 1971. *Modernization and Diversity in Soviet Education.* New York: Praeger.
- Programma. 1990. = Programma "nacional'nye (nerusskie) učebnye zavedeniia RSFSR (1990-1995 gg.)." Moscow: Ministerstvo narodnogo obrazovanija RFSSR.
- Rekomendacii. 1989. = "Rekomendacii respublikanskoj naučno-praktičeskoj konferencii 'Aktual'nye problemy razvitija, izučeniia, prepodavaniia marijskogo jazyka

- i literatury v uslovijax marijsko-russkogo dvujazyčija'." Aktual'nye problemy literaturnogo jazyka. Joškar-Ola: Marijskij NII.
- Rezoljucija. 1990 = Rezoljucija 1-go Vsesojuznogo s"ezda obščestva nacional'nogo vozroždenija 'Mastorava'. Mokša 4.
- Sabatkov, R. B. 1989. "Razvitie metodik rodnogo, russkogo jazykov i literatur v nacional'noj škole za 70 let sovetsoj vlasti i zadači, stojaščie pered NII nacional'nyx škol v uslovijax škol'noj reformy." Problemy dvujazyčija v nacional'noj škole. Iževsk: Udmurtija.
- Šaljaev, N. S. 1977. "K voprosu o jazykovyx processax v sel'skoj mestnosti Mordovii." Trudy Mordovskogo NII, 54.
- Sanukov, K. 1990. "Mariji synymutyn šem kečylaže." Ončyko 6.
- _____. 1991. "Mareillekin paikka uudistuvassa maailmassa." Maailma ja me 9.
- Sanukov, K. N. 1992. Marijcy: prošloe, nastojaščee, buduščee. Joškar-Ola: Marij Ušem.
- Simon, G. 1986. Nationalismus und Nationalitätenpolitik in der Sowjetunion. Baden-Baden: Nomos Verlagsgesellschaft.
- Sovetkin, F. F. 1958. "Rodnoj jazyk v nacional'noj škole." Natsional'nye školy RSFSR za 40 let. Moscow: Izd-vo Akademii pedagogičeskix nauk.
- Stepanov, A. F. 1989. Prosvetitel' S. A. Nurminskij. Joškar-Ola: Marijskoe kn. izd-vo.
- Sukunov, X. X. 1989. "Voprosy dal'nejšego razvitija nacional'no-russkogo dvujazyčija v uslovijax reformy školy." Problemy dvujazyčija v nacional'noj škole. Iževsk: Udmurtija.
- Taldin, N. V. 1958. "Mordovskaja škola za 40 let Sovetskoj vlasti." Nacional'nye školy RSFSR za 40 let. Moscow: Izd-vo Akademii pedagogičeskix nauk.
- Veden'kin, A. 1990. "Aralena mo jylme pamašnam?" Ončyko 5.
- Vedomosti. 1959. = Vedomosti Verxovnogo Soveta RSFSR 16.
- Vedomosti. 1990. = Vedomosti S"ezda narodnyx deputatov SSSR i Verxovnogo Soveta SSSR 19.
- Volkov, G. N. 1989. "Osnovnye puti razvitija nacional'noj školy i ee sovremennye problemy." Problemy dvujazyčija v nacional'noj škole. Iževsk: Udmurtija.
- VS. 1990. = Vestnik statistiki 10.
- Xuzangaj, A. 1988. "Pravo na nasledstvo." Družba narodov 9.

APPENDIX

School curriculum plan in native language, literature and Russian in national schools of the RSFSR: 1989-90 school year, weekly amount of hours by grade and type of school:

| Type of school and subject | The number of weekly hours by grade: | | | | | | | | | | | | | |
|--------------------------------|--------------------------------------|----|----|----|------------------|----|----|----|----|----|----|----|-----|-----|
| | 1. | 2. | 3. | 4. | 1. | 2. | 3. | 5. | 6. | 7. | 8. | 9. | 10. | 11. |
| | (11-year school) | | | | (10-year school) | | | | | | | | | |
| Type (1) school | | | | | | | | | | | | | | |
| Native language and literature | 6/4 | 5 | 5 | 5 | 8/6 | 5 | 5 | 6 | 4 | 3 | 4 | 3 | 2 | 2 |
| Russian language | 4/6 | 8 | 10 | 9 | 6/8 | 10 | 9 | 6 | 6 | 5 | 3 | 2 | 1 | 1 |
| Literature | - | - | - | - | - | - | - | 3 | 3 | 2 | 2 | 3 | 4 | 3 |
| Type (2) school | | | | | | | | | | | | | | |
| Native language and literature | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 |
| Russian language | 9 | 10 | 12 | 11 | 12 | 11 | 10 | 7 | 6 | 4 | 3 | 2 | 1 | 1 |
| Literature | - | - | - | - | - | - | - | 4 | 3 | 2 | 2 | 3 | 4 | 3 |

'SPLIT LOCATIVITY' IN OLD GEORGIAN AND SVAN

H. Paul Manning
University of Chicago

Introduction. Both Old Georgian and Svan (Upper Bal dialect) display a differential marking pattern in the 'local cases' that is determined by independent variables of locative role type and inherent denotational content of Noun Phrases (what has been called 'animacy') (Šanije 1976, K'ek'elije (1956), Vogt 1988 [1947] (for Old Georgian); Abesadze 1984 (for Svan).) In this paper I will attempt to explicitly characterize these 'Split Locative' systems in terms of an independently developed theory of morphosyntax (Silverstein 1976) which has proven to be able to characterize other 'split marking systems' in the grammatical cases to a fine degree of delicacy. In particular, I will argue that the boundary for split marking in Old Georgian is [\pm Human], for Svan [\pm Kin]. I will also attempt to give the rudimentary essentials for a typology of such 'split locative systems.' Lastly, I will attempt to clarify certain problems with the use of the genitive in some of these locative splits: I will argue that, *pace* Šanije (1976), Abesadze (1984), the genitive in these systems may play a pseudo-derivational role with no semantic value ('rank-shifting'), rather than form an 'elliptical' expression with a null N head.

Theoretical Preliminaries: Case Marking 'Splits.' The distribution of nominative-accusative versus ergative-absolutive case-marking within the same language ('split ergativity') can be demonstrated to be a regular consequence of interactions between independent variables, minimally of predicate role (agent, patient, and so forth) and the lexically coded denotational content of NPs ('animacy') (see, for instance, Silverstein 1976). The independent variable of 'inherent denotational content' is defined only with respect to a 'space' of features of NP categories in which *continuous regions* are defined and characterizable by the plus or minus value of some feature(s). This has sometimes been called the 'animacy hierarchy' (as if a one-dimensional space), or, as I will call it, following Silverstein 1976 'the reference space.'¹ For each of these bounded regions of the

¹The fact that the features that constitute the 'reference space' are in themselves assymmetric (i.e., [+Kin] presupposes [+Human], but not the reverse) tends to create the sensation of a unilinear 'hierarchy' ordered by some single feature ('animacy') (see Silverstein 1982:229). Generally speaking, my reference to a 'feature-space' or 'reference-space' may be taken to be the same as the 'animacy hierarchy' where no misunderstanding results.

reference space, a particular differential formal mapping ('case-marking') between propositional role and NP type is derived as a *dependent variable*. Note that by 'case-marking' I will generally mean any difference in formal marking:

What has become clear about CASE-MARKING systems---however the FORMAL indication is made...is that the reference space is relevant to any particular language in a very specific way: it is divided into subspaces at cut-off parameters that form distinct, but---and this is crucial---continuous regions bounded at the plus- vs. minus-values of some feature or features. Within each of these subspaces, a particular mapping occurs between predicate role ("thematic relation") and overt CASE-marking for the basic, grammatical CASEs of central propositional value. (Silverstein 1985:17)

The well-documented case-marking asymmetries in the 'basic, grammatical cases' (the 'core arguments' such as subject and direct object) gives us a framework for similar asymmetries in the so-called 'local cases' ('peripheral arguments' such as are not governed by the predicate). I will use the term *split Locativity* to refer to locative constructions (that is, constructions encoding peripheral 'locative' roles) that show a non-null interaction between the *independent variables* of sentential role and NP 'animacy.' Since the framework I am using (that of Silverstein 1976) was devised for marking asymmetries in the 'grammatical cases,' some modification is in order. I will take case-marking in locatives to be a *dependent variable* resulting from the interaction of the following *independent variables*:

- (1.) The inherent referential content of the NP that bears the role.
- (2.) The sentential or propositional role which the NP is to bear.

A case-marking is simply a differential morphosyntactic marking that results from the interaction of at least the above (and possibly other) independent variables. For each case-marking pattern we may specify a mapping which states (1.) the syntactico-semantic DOMAIN to which the rule applies, (2.) the specific sentential or propositional ROLE, (3.) a continuous REGION (subspace) of the reference space, bounded at the PLUS or MINUS values of some feature, for which we may say that this case-marking pattern holds, whether or not any others do, and (4.) what the specific case-marking is, formally. In practice, for some single role, I will express all the marking patterns which hold, usually in the form of a visual representation. The rule will be of the form:²

²There are unfortunate ambiguities in many of the terms I will be using, between names of role types (and thereby 'functions' of a given morphological case), the traditional names of morphological cases, specific case-marking patterns, and feature names of inherent denotational content. In order to eliminate some of this

Locative hierarchy: $F_{x-n}, \dots, +F_x / -F_x, \dots, F_{x+m}$, where $F_x = [\text{Feature}]$

i. above $-F_x$, assign Case A.

ii. below $+F_x$, assign Case B.

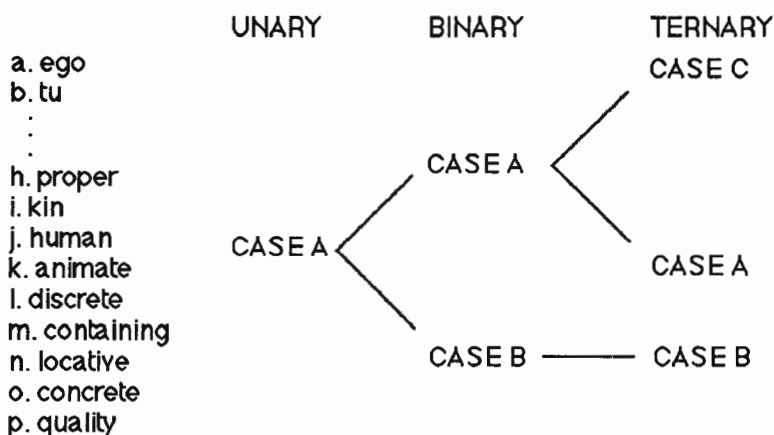
Domain: all objects of adpositions.

Split locative systems. These marking asymmetries in the so-called 'local cases' (peripheral arguments encoding 'locative' roles), what I will call 'split locative systems' (which is intended to remind us of other 'splits' in case marking systems, such as 'split ergativity' and so forth) are really rather widespread typologically, but as with such asymmetric case-markings in the so-called 'grammatical cases,' considerable variation exists. First of all, case marking systems of the grammatical sort have been found to differ in terms of the *number* of bounded regions of the reference space that produce a different case marking when interacting with a single sentential role (Silverstein 1976). The number of such continuous bounded regions of the reference space allows us to define a case system as being, *for some role*, unary, binary or ternary. A *unary system* treats the entire reference space as being *a single bounded region*, and there is exactly one differential formal mapping (let us call this CASE A) between a role and all NPs.³ In effect, all NPs are treated exactly alike. A *binary system* will have *two differential formal mappings* (CASE A and CASE B), each of which codes a mapping between the role and some region of the hierarchy, such that between them there is at least one differential case-marking available for any NP that bears that role, and the two mappings do not both include all regions in the reference space. A *ternary system* will have *three such mappings*, the cut-off point for each being set somewhere in the middle

ambiguity, NP features will be generally enclosed in square brackets ([+F]). I will use the term 'case-marking (pattern/system)' to refer to a system of differential markings that hold for some role type, with 'CASE A,B,C' to refer to specific formally differentiable mappings that are terms in such a system. Alas, morphological cases are often named after the roles or functions with which they are most frequently associated. Rather than have recourse to orthographic subtleties, I will attempt to eliminate this ambiguity discursively, referring to 'instrumental function' or 'role,' and 'instrumental case,' for instance. Terms derived from the literature on thematic roles (such as 'source,' 'goal') will serve as notional characterizations of specific roles, and never morphological cases. Unfortunately, it will sometimes be convenient to refer to a specific morphological case in a specific function (marking a specific role) more telegraphically. Therefore, I will (with apologies) use hyphenated forms like 'instrumental-ablative,' meaning 'the use of the instrumental case in the ablative function.'

³My usage here differs slightly from that of Silverstein 1976. In his usage, '*n*-ary' refers to the number of different case-markings, and '*n*-way' for the number of significant bounded regions of the reference space.

of the reference space and spreading from the top or the bottom of the space. Call the third CASE C. The predicted relative dispositions of the various CASEs (A, B, C) in the reference space (here simplified), and their hierarchical markedness relations, is depicted in the following table:



This table embodies the hypothesis that split locative systems (of whatever elaboration) have recurrent structural properties. A ternary system is composed of two interlocking binary oppositions, rather than a single ternary opposition. A simple binary system will always oppose a residual case (CASE A) to a superimposed case of restricted distribution centering on the feature [+Locative] (CASE B). Note that the relationship between CASE A and CASE B may or may not be disjunctive, this will have to be stipulated by the rule:

Binary Locative hierarchy: $F_{x-n}, \dots, +F_x / -F_x, \dots, F_{x+m}$,
 where F_x = [Feature], and $+F_x > [+Locative]$

1. Below $+F_x$, assign CASE B.

2. Assign CASE A residually.

Domain: e.g. locative, allative *vel sim.*

A ternary system superimposes a second binary opposition between CASE A as *inherited from the first binary opposition* and a new case-marking of restricted distribution (CASE C) spreading from the opposite direction of CASE B. Structurally, CASEs A and C are both opposed to CASE B, and it is thus proposed that CASE A and CASE C will share structural features not shared by CASE B. If the opposition between CASE A and CASE B involves the presence or absence of

adpositions, and CASE A has adpositions, then both CASE A and CASE C will be unified as 'adpositional' case marking as opposed to CASE B. CASE A, it is predicted, will spread from the bottom of the hierarchy, and will therefore not be disjunctive with CASE B, but will be disjunctive with CASE C.

Ternary Locative hierarchy: $F_{x-n}, \dots, +F_x / -F_x, \dots, F_{x+m}$,

where $F_x = [\text{Feature}]$, and $F_x > [+Locative]$

1. Above $-F_x$, assign CASE C.
2. Below $+F_x$, assign CASE A.

Latin will provide a convenient illustration of a binary system. Locative relations in Latin are generally expressed by the use of some adposition which in turn is said to govern a specific case. The specific value of the preposition may depend, as in Classical Greek, on the case it governs. We will call this case-marking pattern CASE A. Thus:

Source: *ab, dē, ex* + ablative case

Destination: *ad, in* + accusative case

Location: *in* + ablative case

However, as nearly every student of Latin will know, 'With names of towns and small islands, and with *domus* and *rus*' (Allen and Greenough 1903:269) a different case-marking pattern (CASE B) prevails, to wit:

Source: ablative case

Destination: accusative case

Location: locative case

Such systems, where CASE A is residual and CASE B is limited to $[+Locative]$ nominals, are extremely widespread, although the pattern is seldom as clean as that in Latin. The formal difference of Latin is one found in many languages as well, between an ungoverned case and a case governed by a preposition: an analogous pattern is found even in English 'bare NP adverbials' (Larson 1985).⁴

⁴Marking splits at the lower ends of the reference space have been pointed out before (see, for instance Dixon's discussion (1976) of the distribution of the dative and locative cases in Yidinj). There is thus nothing particularly innovative about my application of the referential hierarchy to this end of the reference space. A full justification for the entire reference space (especially the lower 'inanimate' end) and the principles according to which it is organized can be found in Silverstein (1987). To summarize, in his earlier formulation (1981:242) the organizing principle of the reference space is 'the unavoidability and transparency of metapragmatic reference' (a formulation that I will not attempt to

A binary locative split by definition divides the reference space into two (quite possibly overlapping) regions. *Ex hypothesi*, in such a system, 'CASE B' will form a region containing [+Locative], while the other, CASE A, is residual. We may take a binary pattern to represent a superimposition of a specialized case marking (CASE B) on a simple unary system (CASE A), and this specialized marking will always spread from [+Locative].

This is a vacuous formulation as it stands, unless we have some independent means of 'recognizing' each CASE *formally*. The following formal criteria may be suggested:

(1.) *Distributional Residualness*: In any binary split locative system, CASE B will be limited to a subset of the total reference space containing minimally [+Locative]. CASE A will be residual. Thus, all other things being equal, CASE A's domain may contain that of CASE B by inclusion. Thus, Yidinj (as described by Dixon 1976) has a binary split locative system featuring two morphological cases: dative and locative. Locative is limited to [-Human] nominals, while dative is residually possible throughout the reference space. This is a limiting case of complete overlap: dative case may be used with any NP, locative is restricted to [-Human] NPs. Thus, dative is our CASE A, locative our CASE B.

(2.) *Structural Markedness*: In a binary split locative system, CASE A may more structurally complex than CASE B, not vice versa. Thus, if a split system manifests a contrast between adposition versus no adposition, CASE A will feature adpositions, CASE B will not. Note that the above definition does not require

unpack here.) However, as Silverstein (1987, personal communication) makes clear, as we move down the reference space from the speech act indexicals, at a certain point the heads of NPs do not refer all by themselves, but rather provide characterizability conditions on potential reference. We move from unavoidable metapragmatic reference to simple sense categories. In fact, in Silverstein 1987 we find that four different properties (or 'components of characterizability conditions') are involved, which have a 'differential prominence...in the different regions of the hierarchy.' These are: (1.) 'the indexical conditions on felicitous usage in the instance.' (2.) 'The socio-historical warrant for usage.' (3.) 'The structure of (shared) extendables.' (4.) 'The specific morphosyntactic structure of the language as it corresponds to the grammatical categories in the linguistic expression at issue.' As we move down the space, from speech act indexicals though proper names to animates and finally spatial, the differential prominence of each of these 'components of characterizability conditions' changes in the same order. The lower ends of the hierarchy are less indexical of the speech act (1.) than they are of the morphosyntactic structure of the language as a whole (4.). I cannot hope to fully justify the lower ends of the reference space here, I wish merely to point a *prima facie* relevance of the lower ends, as well as the higher ends, to the descriptive problem at hand. Those interested in the theoretical grounding of the reference space are advised to consult the above article.

CASE A to exhibit greater structural markedness than CASE B, as it does not in Yidinj:

| | CASE A | CASE B |
|---------|-------------------|---------------------------------|
| Latin | prep+case | case (Allen and Greenough 1903) |
| Yucatec | prep | Ø (Hanks 1990) |
| English | prep | Ø (Larson 1985) |
| Yidinj | case ₁ | case ₂ (Dixon 1976) |

(3.) *Semantic Elaboration*: Insofar as the formal difference between CASE A and CASE B is one of difference of structural elaboration, and the structural element present in CASE A but not in CASE B (for instance, an adposition) is a member of a paradigm of similar elements with computible sense value, we may say that CASE A is semantically elaborated vis a vis CASE B. Typically, CASE B will be limited to a few very general locative roles (maximally location, destination, source, minimally location or destination), while CASE A may express a wide variety of finely differentiated locative roles. Note that this need not always be the case: Yucatec Mayan (Hanks 1990) is typical of Mayan languages in having a binary system opposing bare NP adverbials to a prepositional system of exactly one term. Clearly CASE A may not make any separate contribution to sense in such a situation.

A ternary marking system seemingly always superimposes a third case-marking contrast on a system with all of the formal properties of a binary system. The third CASE, CASE C, is always formally a subtype of CASE A. Formally, CASE C will have the following properties:

(1.) *Distributional Residualness*: CASE C will always spread (implicational) from the top of the reference space down. Thus, it will spread from the opposite direction of CASE B. CASE C will generally complement CASE A, which in a ternary system generally overlaps with CASE B.

(2.) *Structural Markedness*: CASE C will generally be a formal variant of CASE A: for instance, in a ternary system, both CASE A and C may feature adpositional marking over and against CASE B. However, CASE C will be the structurally more elaborate of the two if there is any difference in structural markedness.

(3.) *Semantic Elaboration*: Insofar as CASE A and C are formal variants of one another (in that if CASE A uses adpositions, so too will CASE C), CASE C would be expected to be able to express the same array of locative relations as CASE A. However, CASE C may involve a neutralization of locative oppositions as part of its expression. Frequently, inessives and ablatives are neutralized in CASE C to adessives and elatives, the notion of 'interiority' being neutralized.

The above are stated somewhat optimistically as testable hypotheses, not as proven typological universals.⁵ They are intended to provide a necessary context and some handy monickers for the examination of two ternary split locative systems: those found in Old Georgian and Svan.

Old Georgian Split Locatives. Old Georgian displays two split locative patterns. The first, described by Vogt (1988 [1947]) and Šanije (1976), opposes ungoverned local uses of the instrumental (in an 'ablative' function) and adverbial cases (in the 'lative' function) with prepositionally governed local uses of those cases. The second, described by Šanije (1976) and more fully by K'ek'elije (1956)⁶, involves a small set of adpositions (*zeda*, *tana*, *činaše*) which govern either the dative or genitive cases depending on the inherent denotational content of the NP governed.

The instrumental case in Old Georgian has three broadly discernible functions: marking the instrumental, comitative (with *-urt*), and ablative (with or without the adposition *-gan*) (Vogt 1988[1947]) roles.⁷ The pure instrumental function (hence the name of the case) might be taken as basic, as it is only in this use that both the long and short variants of the morphological case form (*-it*, *-ita*) commute:

sixarulit μετα χαράς

Joy-INSTR

'With joy.'

sixarulita mit ' από τῆς χαρᾶς

Joy-LongINSTR That-INSTR

'From the joy.'

The conditions governing the choice in Old Georgian between long and short case forms has been adequately discussed by Vogt

⁵ As Silverstein (personal communication) points out, the above principles are all jointly applicable aspects of the structure of telescoping case-markings, such that a maximally differentiated system of differential case-markings found somewhere in the language (CASE A) is 'telescoped' into a minimally differentiated system found elsewhere (e.g. CASE B).

⁶ I thank Winfried Boeder for bringing this article to my attention.

⁷ Once again, the ambiguity here between the name of the morphological case ('instrumental') and its various functions ('instrumental,' 'comitative,' etc.) is unfortunate.

(1988[1947]). Suffice to say here that the exponence of the case-marking is cumulative with that of what we might loosely call 'determinateness.'

The comitative function of the instrumental case involves (1.) only the short variant (*-it*) of the instrumental affix and (2.) always cooccurs with the adposition *-urt* (examples from Vogt 1988[1947]:212).

iaḡob-it-urt da iohan-it-urt Mk. 1.29 μετὰ Ἰακώβου καὶ Ἰωάννου
Jacob-INST-with and John-INST-with
'With Jacob and John.'

The ablative function (marking 'source of motion' role) of the instrumental, like the comitative, is restricted to the short variant, but may or may not be governed by the adposition *-gan*. The ungoverned instrumental-ablative is limited primarily to [+Locative] nominals, and the governed instrumental with *-gan* is residual:

aḡmosaval-it movides Mt.1.1
East-INSTR Came-3Pl-Aor
'They came from the east.'

dasabam-it-gan (Šanije 1976:143)
Beginning-INSTR-from
'From the beginning.'

According to Šanije (1976:44) Vogt (1988[1947]), the (governed or ungoverned) instrumental (in 'ablative' function) may not occur with proper names. However, it appears from the data (see below) that any nouns with human denotata are similarly banned. As the use of the instrumental (in 'comitative' function) above with precisely such nouns indicates, it is not the instrumental *qua* morphological category which is so banned (that is, proper names do not have a 'defective paradigm'), rather, it is the instrumental case in this specific function ('instrumental-ablative'). However, we do find examples where the instrumental-ablative (with *-gan*) is added to a *derived theme of the genitive* (as both Šanije and Vogt note):

abraham-is-it-gan Mt.1.17 ἀπὸ Ἀβραάμ
Abraham-GEN-INSTR-from
'From Abraham'

Passing now from the instrumental-ablative to the adverbial-allative,

we find a similar situation. The adverbial case may occur as an ungoverned case marking lative (destination of verbs of motion) roles with [+Locative] nominals:

movides batpaged Mt.21.1 ἦλθον εἰς βηθφαγή

Come-3rdPL-Aor B.-ADV

'They came to Bethphage.'

However, the adverbial case *qua* morphological category is peculiar in that it lacks the short-long opposition typical of other oblique cases (dative (-s, -sa), instrumental (-it, -ita), genitive (-is, -isa)) in Old Georgian. Indeed, the semantic opposition ('determination') coded by the selection of short versus long case affixes is neutralized in the adverbial. Consequently, when 'determination' must be encoded cumulatively with the adverbial-lative, the adverbial is neutralized to a (long) dative in -sa (see Vogt 1988[1947]):

[-Determined]: adverbial [+Determined]: long dative in -sa

mta-d Mk. 13.4 εἰς τὸ ὄρος

'To the mountain.'

mta-sa maḡal-sa Mk. 9.2 εἰς ὄρος

ὑψηλόν

'To the high mountain.'

Nominals which are not [+Locative] require the adverbial with the postposition -mde. But it appears that even [+Locative] Nominals may also occur with -mde (indeed, these are the only examples Šanije provides (1976:144)):

joḡoxet-ad-mde

Hell-ADV-towards

'To Hell.'

c-ad-mde

Heaven-ADV-towards

'To Heaven.'

Lastly, as with the instrumental-ablative, the adverbial-lative may occur with proper names, and, it appears, with [+Human] nominals generally, only if the adverbial (in its variant -a) is added to a derived theme in the genitive, in a manner exactly parallel to the instrumental-ablative above (as Vogt notes). Instances of apparent Genitive case with -mde listed as such in Šanije (1976:143) have been plausibly explained by Vogt (1988 [1947]) as a genitive theme and -a affix of

the Adverbial:

(1.) [+Human] NPs:

mep-is-a-mde
King-GEN-ADV-towards 'To the king.'

(2.) [+Proper] NPs:

davit-is-a-mde Mt.1.17 ἕως Δαυείδ
David-GEN-ADV-towards 'Up to David.'

(3.) Third person pronouns with human denotata:

čem-da-mde John 21.23
me [GEN]-ADV-towards
'Towards me.'

Thus, we have a ternary system of case marking (where ABLATIVE and LATIVE refer to 'source' and 'goal' roles respectively):

| | ABLATIVE | LATIVE |
|--------------------|--------------------------------|--------------------------|
| CASE B [+Locative] | instrumental | adverbial ~ dative (-sa) |
| CASE A | instrumental + -gan | adverbial + mde |
| CASE C [+Human] | genitive derived theme+ CASE A | |

Binary Locative hierarchy: $F_{m-n}, \dots, +F_m / -F_m, \dots, F_{m+m}$, where F_m = [Containing]
Below $+F_m$, assign Case B. [-adpositional]
Assign Case A residually. [+adpositional]
Domain: Ablative and Lative

Ternary Locative hierarchy: $F_{j-n}, \dots, +F_j / -F_j, \dots, F_{j+m}$, where F_j = [Human]
Below $+F_j$, assign CASE A.
Above $-F_j$, assign CASE C.
Domain: Ablative and Lative

In addition to the above ternary split locative system, Old Georgian has another binary split locative system. Three adpositions (*tana* 'with,' *zeda* 'on' and *činaše* 'before') govern either the dative or genitive case according to the inherent denotational content of the governed

nominal.⁸ K'ek'elije (1956) has done an exhaustive survey of the distribution of these two cases vis à vis the inherent denotational content of NPs they occur with and came to the conclusion that the genitive case was limited to use on [+Human] NPs (*gonieri* 'intelligent'), the dative to [-Human] NPs (*aragonieri* 'unintelligent'). As K'ek'elije notes, this classification is also that which determines choice of possessive verbs (*hqavs* 'to have (someone)'—*akvs* 'to have (something)') and the distribution of interrogative pronouns (*vin* 'who'—*ra* 'what')). We may note further that this same boundary figures in the above ternary system in Old Georgian.

Locative hierarchy: $F_{j-n}, \dots, +F_j/-F_j, \dots, F_{j+m}$, where $F_j = [\text{Human}]$

Below $+F_j$, assign Dative.

Above $-F_j$, assign Genitive.

Domain: [___ P] where P is a member of {zeda, tana, činaše}.

Distribution of Genitive and Dative Cases in Old Georgian Split Locatives

⁸Note that these three adpositions are all 'locative' as opposed to 'directional' (like *-mde* 'towards,' *-gan* 'from'). These form, in fact, the basic set of 'locative' (coding some 'location' roles) adpositions. There is thus some semantic coherence to this class (since only *šina*, *šida* 'in' is lacking from this alternation). The latter is banned for pragmatic reasons from human denotata in normal usage. Further note that there is a parallelism across the board between Case A, Case B, and Case C codings in each locative role type:

| Role type | Case B | Case A | Case C |
|-----------|----------------------|---|---|
| | [+ locative] | [residual] | [+ human] |
| Locative | dat | dat + <i>tana</i> , <i>zeda</i> , <i>činaše</i> | gen + <i>tana</i> , <i>zeda</i> , <i>činaše</i> |
| Lative | dat/adv | advl + <i>-mde</i> | [[[gen] adv] <i>-mde</i>] |
| Ablative | instr | instr + <i>-gan</i> | [[[gen] instr] <i>-gan</i>] |
| | [Case ₁] | [Case ₁ + P] | [[[gen] Case ₁] P] |

The locative differs from the directional cases in that it replaces the dative with the genitive rather than adding the normally governed case outside the genitive in Case C.

| | | |
|---------------|--|----------|
| a. ego | | |
| b. tu | | |
| ⋮ | | |
| h. proper | | GENITIVE |
| i. kin | | |
| j. human | | |
| k. animate | | |
| l. discrete | | |
| m. containing | | DATIVE |
| n. locative | | |
| o. concrete | | |
| p. quality | | |

I will give examples only for the adposition *zeda* 'on,' for which the documentation (all from K'ek'elije 1956) is more complete:

(1.) The dative case is assigned to [-Human] NPs: nouns denoting (a.) inanimates, (b.) animates, and (c.) pronouns substituting for these (K'ek'elije 1956:306-308). Note that pronouns do not encode categories of animacy or gender lexically.

| (a.) [-Animate] NPs | (b.) [-Human] NPs | (c.) [-Human] Pronouns |
|---------------------|-------------------|------------------------|
| gza-sa zeda | ḡor-sa zeda | mas zeda |
| road-DAT on | mule-DAT on | it-DAT on |
| 'on the road' | 'on the mule' | 'on it.' |

(2.) The genitive case is assigned to [+Human] NPs: (a.) proper names of humans, (b.) human kin terms, (c.) human status terms and other nouns denoting humans, and (d.) pronouns denoting humans (K'ek'elije 1956):

| (a.) [+Proper] NPs | (b.) [+Kin] NPs | (c.) [+Human] NPs | (d.) [+Human] pronouns |
|--------------------|-----------------|-------------------|------------------------|
| iesu-ys zeda | mam-isa zeda | mep-isa zeda | mis zeda |
| Jesus-GEN on | Father-GEN on | king-GEN on | him/her-GEN on |
| 'on Jesus' | 'on father' | 'on the king' | 'on her' |

(3.) There is, however, a residue of items which display quixotic behavior. Consider the case of *eri* 'people, nation, crowd,' which appears in the dative regularly with *zeda* and *tana*, but in the genitive with *činaše*: I can give no explanation for the overlap between categories

here: however, not that the two postpositions of the class (zeda and tana) both govern the dative in this instance, and the lone preposition (činaše) governs the genitive:

| | | |
|---------------|-----------------|--------------------|
| ersa zeda | ersa tana | činaše erisa |
| crowd-DAT on | crowd-DAT with | before crowd-GEN |
| 'on the road' | with the crowd' | 'before the crowd' |

Svan Split Locatives. Svan (Upper Bal dialect) displays a ternary split locative system of much greater pervasiveness than Old Georgian. The following discussion is based on that of Abesadze (1984), but I have also drawn data from text collections provided by Gudjedjiani and Palmaitis (1986) (henceforth G&P 1986) and Šaniĵe et al. 1939). Like Old Georgian and Modern Georgian, certain ungoverned morphological cases may mark location and perhaps some other relations:

muš-s liz lequnzela gun (G&P 1986:38)
 Mush-DAT is vineyard very
 'In Mush there are very many vineyards.'

Such ungoverned uses conform to the characterization of CASE B in a binary system: they are restricted distributionally to a subregion of the reference space that contains [+Locative] NPs, they are structurally less complex than other locatives (CASE A and CASE C); and they are semantically less elaborate than other locatives in the array of different Locative roles they can mark.

Superimposed on this is a rather more elaborate bifurcation of the reference space at the plus or minus value of the feature [\pm Kin], whereby all adpositions assign genitive case to NPs that are [+Kin], and Dative case to those that are [-Kin]. I will take the dative case assignment to be residual in this instance (see below). The above rule only applies to those postpositions that assign dative case residually.

Postpositions in Svan (Upper Bal dialect) come in two general varieties: postpositions which generally cliticize to the lexical head of the NP they govern and assign dative case (example 1) , and postpositions which do not cliticize and assign genitive case (example 2). Each class can be defined syntactically, and unlike in Old Georgian, there appears to be virtually no lexically variable government of case. The

first class appears to include all basic postpositions and forms a closed class, the second appears to be derived and has a nominal character. I will refer to the first class as dative postpositions, and the second class as genitive postpositions. Note that case affixes in Svan undergo truncation in a manner similar to Modern Georgian. This phenomenon will be discussed below.

(1.) lemasgw-šāl (Holy Fire 15 (Šanije et al. 1939 #55:53))
fire [Dative]-like 'Like fire.'

(2) zagr-iš i zagr-i nēsga (G&P:51) 'Between a ridge and a ridge.'
ridge-Gen and ridge-Sh. Gen between

Dative postpositions, not genitive postpositions, display apparent variable government of case according to the inherent denotational content of the NP governed (See Abesadze 1984 for a full discussion). The system is binary in the sense that where postpositions occur, there are two relevant subregions of the reference space which determine differential marking. Genitive case is assigned to NPs that are [+Kin] and above (as well as pronouns with human referents), dative case is assigned to NPs that are [-Kin] and below. In the following examples I will suppress the sentential context for the postpositions in question, since it is not immediately relevant. The full context can be found in the sources cited.

(1.) Genitive case is assigned to (a.) first and second person pronouns, which assume their possessive form, (b.) third person pronouns when they have human denotata, (c.) proper names of humans, (d.) kin terms (see Abesadze 1984:9):

(a.) First and Second person Pronouns:

nišgwēy-caxān (G&P 1986:120)
our(Excl)[Poss]-at
'Unto us'

mišgwa-te (Abesadze 1984:9)
my[Poss]-DAT-to
'To me'

(b.) Third person pronouns with [+Human] denotata:

mineš-te (Dogs Saint George¹⁴ (Šanije et al. 1939 #106:90-91))

They[GEN]-to

'To themselves'

mičēš-xān (Mourning 9 (Šanije et al. 1939 #44:44))

She[GEN]-from

'From her (family)'

(c.) [+Proper] NPs:

jantemaz-iš-te (Sosruq 67 (Šanije et al. 1939 #184:162-3))

Jantemaz-GEN-to

'To Jantemaz.'

(d.) [+Kin] NPs:

mine mēš-te (Abesadze 1984:9)

Their-ShGen Father-GEN-to

'To their father.'

(2.) Dative assignment is residual, but it is crucial to note that (a.) NPs with human denotata, (b.) third person pronouns with non-human denotata, and (c.) proper place names all occur in the dative. Abesadze (1984) appears to claim that all NPs with human referents may receive genitive marking, but that this is impossible for non-human referents. She treats these examples as an example of 'wavering.' It seems to me that it is nothing of the sort: [-Kin] NPs (including [+Human] NPs) simply always appear in the dative, with the usual exceptions for third person pronouns⁹ and [+Kin] NPs always appear in the genitive:

(a.) [+Human] NPs:

tebuš-er-məq (Lamproba 32 (Šanije et al. 1939 #13:15))

Tebush.-PL-DAT- at 'At the Tebush's.'

xälx-isga (G&P 1986:117)

People-DAT-in 'Among the people.'

⁹ Cf. Silverstein 1985:20: 'An anaphoric form seems always to be treated at least as "good" and "Agent-of" a predicate, and at least as "bad" a "patient-of" a predicate, as its corresponding antecedent, all other things being equal.' The same applies in split locative systems. For further discussion see Silverstein 1987:141ff.

mezwbēl-ār-məq̄ (Abesadze 1984:12)
Neighbour-PL-DAT-at 'At the neighbor's'

bepšw-caxān (Abesadze 1984:11)
child-DAT-with 'With the child.'

(b.) Third person pronouns with non-human denotata:

eč-xānka (G&P 1986:116)
That-DAT-out-of 'From there'

(c.) Proper place names:

šwān-te (Liturgy Archangel 8 (Šanije et al. 1939: 14:16))
Svanetia-DAT-to 'To Svanetia.'

Sufficient examples exist of dative usages, most particularly in the [+Human] category, to demonstrate that [\pm Kin], and not [\pm Human], is the boundary between the two relevant subregions of the reference space. Thus, it appears that Old Georgian differs subtly from Svan in this respect, though it is hard to draw hard conclusions in the light of the data.

Svan Locative hierarchy: $F_{i,n}, \dots, +F_i / -F_i, \dots, F_{i+m}$, where $F_i = [\text{Kin}]$

Below $+F_i$, assign Dative.

Above $-F_i$, assign Genitive.

Domain: [P], i.e. all 'Dative adpositions.'

Domain of Genitive and Dative Cases in Svan Split Locatives

| | | |
|---------------|--|----------|
| a. ego | | |
| b. tu | | |
| : | | |
| : | | GENITIVE |
| h. proper | | |
| i. kin | | |
| j. human | | |
| k. animate | | |
| l. discrete | | |
| m. containing | | DATIVE |
| n. locative | | |
| o. concrete | | |
| p. quality | | |

Formal Problems with the Genitive Case in Old Georgian and Svan. In the above discussion I have blithely presented the contrast between CASE A and CASE C codings in Old Georgian and Svan in terms of 'differential government of case.' Examination of surface coding contrasts may seem to dictate that this be our 'null hypothesis,' but the role of the genitive case in both of these languages is such that a reexamination of its coding role is warranted. The question that arises is whether the role of the genitive in the above examples is 'governed' in a manner at all parallel to, for instance, the dative. The following section will argue that some instances of the genitive in Old Georgian and Svan are examples of what has been called 'rank shifting' (Halliday 1961, Silverstein 1984). *Rank shifting* is a lexical process that uses the inherent derivational potentiality of the genitive case in these languages to 'create stems suitable for hierarchical specification' (i.e., inflection) (Silverstein 1984:281) from stems, words or phrases that are not, for whatever reason. Teleologically speaking, rank shifting is triggered when a nominal stem which, due to defective morphology or inherent (major or minor) categorial specification, is incumbent to a syntactic position from which it is banned. Rank-shifting creates a new substantive stem that is able to be inflected and has the distributional properties of a generic nominal both syntactically and morphologically. This position is similar to, but not identical with, positions taken by Šanije (1976:44) and Abeladze (1984:10-11) on the function of the genitive in Old Georgian and Svan, respectively.

In Old Georgian, the genitive occurring with the three adpositions (*zeda*, *tana*, *činaše*) commutes with the dative. In this case, there is little or no evidence that we should reject our adopted null hypothesis of direct differential government, there being no evidence of hierarchical subordination of the genitive to the dative. However, in the instance of the contrast between the governed and ungoverned uses of the instrumental and adverbial, we may note the following:

(1.) The contrast between CASE A (the adpositional variant) and CASE B (the adpositionless variant) is one of what we might call density of adumbration (or 'extended exponence') of the roles ablative and lative. CASE A is structurally more marked than CASE B (the markers of CASE B are a proper subset of CASE A), but the additional marker present in CASE A, and not in CASE B, is a further exponent

of the role ablative or lative, further, it is added *externally* to the marker present in CASE B:

| | ABLATIVE | LATIVE |
|---------|----------|----------|
| CASE A: | X-it-gan | X-ad-mde |
| CASE B: | X-it | X-ad |

(2.) Similarly, the contrast between CASE A and CASE C is one of structural markedness: CASE C is structurally more marked than CASE A in that it includes the markers present in CASE A as a proper subset. However, the marker present in CASE C does not adumbrate or expound any locative category whatsoever: the same additional marker (genitive) is present in both the Ablative and the Lative, and moreover it is *internal* to the markers found in CASE A. Thus, those elements having CASE C coding (indicated by Y in the following diagram) are *ipso facto* distributionally equivalent to those requiring CASE A (X in the following diagram). By being placed in the Genitive case, items of class Y are 'rank-shifted' to the distribution of X, and then receive CASE A coding:

| | ABLATIVE | LATIVE |
|---------|-------------------|------------------|
| CASE A: | X-it-gan | X-ad-mde |
| CASE C: | [x Y-is x]-it-gan | [x Y-is x]-a-mde |

Minimally, then, we may say that (1.) the genitive case is transparently *hierarchically subordinated* to the exponents of ablative and lative roles, (2.) the genitive is *not in itself an exponent of these categories* in this construction; and (3,.) the stem (Y) marked with the genitive *has the distributional privileges of X*, which the stem (Y) not marked by the genitive lacks. We may say then that since there is no contrast in this environment between Y and Y-GEN, Y and Y-GEN are semantically noncontrastive: further since Y-GEN has the distribution of X (in this context only), the function of the genitive is to take an input expression Y and return an expression X such that Y may not occur in this environment and X can.

Since Svan displays a striking parallelism with Old Georgian in its distribution of genitive and dative cases with postpositions (as Abesedze (1984:10) notes), we might wish to preserve their paradigmatic contrast in Svan as well as Old Georgian. There is, however, evidence, albeit

not particularly ubiquitously coded, that hierarchical subordination (CASE A: [X-DAT P], CASE C: [[Y-GEN]-DAT P]) rather than differential government (CASE A: [X-DAT P], CASE C: [Y-GEN P]) is the correct analysis (as, to be sure, Abesadze (1984) notes).

First and second person pronouns in Svan are uninflectable. Hence, when they must be used possessively and so bear agreement, they have peculiar forms which (in the singular) can also encode a limited set of agreement features (when not truncated). Thus (comparing the first singular forms with the first plural exclusive (truncatable material is given in parentheses)):

| | Basic | Possessive | Possessive-Dative |
|-------------|-------|------------|-------------------|
| 1 Singular: | mi | mišgw(i) | mišgw(a) |
| 1 Plural: | ni | nišgwey | nišgwey |

The singular series is capable of encoding the dative (and hence oblique) case as a separate stem form, while the plural is not. These forms are also used, as we would expect, when the pronouns are governed by postpositions: they are formally genitives. However, the singular forms seem to use the dative form of the possessive, while in the plural, where the contrast is neutralized, the basic possessive form is used. Hence, there is immediate evidence that the genitive (possessive) is hierarchically subordinated to the dative:

mišgwa-te *mišgwi-te
nišgwey-te

One may now object: if the dative is present added to the derived genitive stem, where then is its formant? Observe that in normal postpositional phrases in Svan, the most ubiquitous, segmentable formant of the dative (-s), in those declensions that have it, is never found when the governed item immediately precedes the governing item. Such 'truncation' occurs more sporadically in Modern Georgian: in Svan, it is the rule with postpositions, and adnominal genitives and agreeing datives which precede the head of the NP (all examples from G&P 1986). Note that the environment for all these rules is one of string adjacency and in inherently directional ('left-looking.') In the following examples, the material between square brackets ('[...]') is truncated.

(1.) Dative -s truncation with postposition.

Trunc.



kor-[s]-te 'To the house.'
House-DAT-to

(2.) Dative -s truncation of agreeing item (appositive prenominal modifier).

Trunc.



ašxw mārā-[s] muswṭi-s (G&P 1986:120)
onc-DAT man-DAT Near-relative-DAT
'One man, a near relative of the deceased.'

(3.) Genitive truncation (and non-truncation).

Trunc.



məžalw-äš i məlx-ä[š] lağlät-te (G&P 1986:117)
Muzhal-GEN and Mulakh-GEN ravage-to
'To ravage Muzhal and Mulakh'

Trunc.



jär-i[š] maxwši säwj-ar-eš (G&P 1986:117)
Army-GEN chief Kabardian-PL-GEN
'The chief of the army of the Kabardians.'

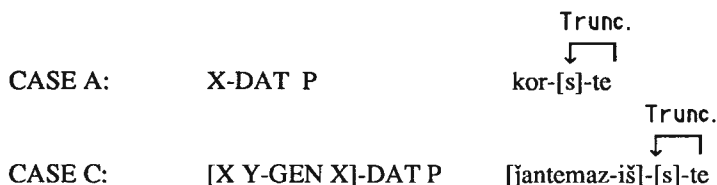
There is, of course, independent evidence that items in this position do carry the dative case: certain declensions (notably the 3rd, 4th, and 5th) have nonsegmentable exponents of the dative (vocalic changes of the stem): these do not truncate:

lemasgw-šāl (see above) 'like fire.'
Fire [Dative]-like

Compare *lemesg* 'fire' (basic Nominative stem).

Indeed, we may ask, given that any governed case seems to be truncated when it immediately precedes the governor, if the genitive is directly governed, why does it not truncate, as it does with those postpositions which govern the genitive across the board? More productively, I would like to suggest that (1.) since, as we will see, all

genitive-derived stems are members of the sixth declension, which has the affix *-s* as its dative exponent, (2.) and this affix is precisely the one which undergoes truncation; and (3.) the dative is in fact assigned to the genitive stem and only those items, like *mišgwi* ~ *mišgwa*, which are capable of representing this dative non-affixally survive to tell the tale. Thus, in Svan there is reasonable evidence, obscured by ubiquitous morphological processes of truncation, that the genitive is hierarchically subordinated to the dative, rather than being directly governed in paradigmatic contrast to the dative (these are essentially the conclusions of Abesadze 1984).



Why then, we may ask, is the genitive case in particular drawn upon to serve this particular function, which is more derivational than inflectional? In order to assess the hypotheses of Šanije on the above construction, as well as a similar proposal of Abesadze's the use of the genitive in Svan, we may wish to delineate the exact range of similar functions of the genitive in both these languages.

Firstly, we may not unjustly compare some uses of the genitive in Old Georgian and Svan to similar uses of the 'substantivizing article' in Classical Greek (Smyth 1920:292-3, 450ff. Goodwin and Gulick 1981:204ff,325ff). Initially we may note that the Kartvelian genitive, like the Ancient Greek article, has two coexistent potentialities or 'moments' (drawing terminology from Haas (1966).) In the first, the *diacritic*, it contrasts *paradigmatically* as a *term* in a *system* of cases: thus by commuting with the dative with three adpositions in Old Georgian, its paradigmatic contrast gives it a diacritic coding value. In the second, the *determinative*, we may surmise, as with the Greek substantivizing article, that anything marked with the genitive is *ipso facto* a nominal.¹⁰ The determinative function of the genitive is simply

¹⁰ In Kartvelian languages, as is well known, both Nouns and Adjectives may rather freely occupy the function of Head of NP (hence N function) or attribute of NP (A function), or indeed predicate: by nominal we will mean 'either N or A.'

that: the fact that whatever it attaches to becomes a *de facto* nominal.¹¹ Both of these potentialities may be said to coexist implicitly in the genitive, and the precise actualization of either will be a matter of its constructional or lexical contexts of use. Note that the genitive may simultaneously serve as a *term* in a paradigmatic contrast (diacritic function) as well as what the Firthians called a *marker* (determinative function: marking (even if redundantly) the constructional role of another item in the syntagmatic plane) (Haas 1966.).

Turning from the diacritic function, we may note that there are at least three uses of the Genitive in Old Georgian and Svan which foreground the determinative function of the Genitive. What these all share is that *a stem or word of whatever categorial specification, by being placed in the Genitive, is returned as a categorial nominal stem of some sort, capable of further hierarchical specification (inflection), due to the 'determinative' value of the genitive.* I will dub these: (1.) elliptical derivation, (2) lexemic (lexical) derivation, and (3) rank shifting.

(1.) The function of the genitive in *elliptical derivation* is to (i.) take an input item (whose category may include N, A, ADV...), (ii.) and create a nominal expression (N ~ A). This expression (iii.) serves as an ad hoc referring expression, (iv.) whose meaning is computable from the input expression by an open function, but the denotation of the input and output are distinct. Thus, it is not derivational in the sense of creating a new lexeme whose sense is concretized when the expression is coined and becomes a member of the lexicon, but rather it derives an 'elliptical' expression that is interpreted generally as if it were an elliptical variant of a full NP. In this usage the Old Georgian genitive is parallel to some uses of the Greek 'substantivizing article,' such that the one can translate the other (using only the long form of the genitive, as Vogt (1988:224 [1947]) notes).¹²

¹¹ As Kevin Tuite (personal communication) notes, other cases can have the effect of rank shifting, especially in delocutionary constructions, in Old Georgian and Svan. He notes that one key difference is that the genitive case leaves the affected stem able to receive further hierarchical specification, whereas the dative does not. As for the range of possible inputs for genitive rank-shifting, the inputs include at least all nominal categories (N and A and their projections), and for some purposes, adverbs and prepositional phrases.

¹² "Les formes longues en -isa sont, au contraire, régulières dans les dérivés

ḡmrt-isa-sa

God-GEN-DAT

τὰ τοῦ Θεοῦ

'The (things) of God'

ḡac-isa-sa

Man-GEN-DAT

τὰ τοῦ ἀνθρώπων

'The (things) of men'

As Šaniġe (1976:42) claims for Old Georgian, and Abesadze (1984:10-11) for Svan, we can assume that our 'derived genitive themes' are genitive adjuncts depending on a null head syntactically (Abesadze's '*verlorengegangenen Determinandum*'), which then receives case according to its syntactic environment.

[_{NP} [N_i-GEN] Ø_N-DAT]

Semantically these elliptical derivations would be open formula of the form:

'The ____ of N_i'

where the open slot can, according to Šaniġe (1976:42), be filled from a limited number of general meanings: 'time,' 'house,' 'place' etc.... Abesadze's analysis appears to follow similar lines. The use of a null head in these expressions appears to be superfluous in a language where 'headship' of NP is defined by position and not by category; I will not pause to critique it. However, the open semantic formula is very slippery indeed. Presumably the meaning would be limited to very few, general candidates, as in the English elliptical expression *at John's*, meaning uniquely 'at John's house' (or similar domicile). Note that in this expression, a sentence like

I went to John's, but he wasn't there.

is perfectly acceptable. If the above 'open semantic formula' is to hold, there must be (1.) no cases where forms not using the genitive nouns comuns, Mc. 16.23 *ara zraxav šen ḡmrtisasa, aramed ḡacisasa* οὐ φρονεῖς τὰ τοῦ Θεοῦ, ἀλλὰ τὰ τοῦ ἀνθρώπων ...οὐ le génitif en -isa ne signifie pas 'de Dieu,' mais 'celui de Dieu,' τὸ θεῖον.' Thus the 'long' variant seems to have a force similar to the substantivizing article here.

have this 'elliptical interpretation' and (2.) those forms using the genitive should hopefully have a rather limited number of general interpretations. The moment we reach a point where we say that 'the area where X is' fits the bill, we are really saying that it encodes nothing. Neither (1.) nor (2.) appear to be true of Svan. The postposition *-mæq*, for instance, seems to mean 'at ___'s residence.' It has this meaning even when it governs the dative (see above). Similarly, in the context of the story from which the prepositional phrase *ǰantemaz-iš-te* ('To Jantemaz') is taken, the motion is simply towards *ǰantemaz* (a human proper name). *Insofar as the distribution of Genitive versus Dative is uniquely determined by interactions between semantic role and noun phrase denotational content, and no contrast is available, we need not assume that the Genitive has semantic moment in one context simply because it does elsewhere.*

(2.) The function of the genitive in lexical derivation is syntactically identical to the above (i.) and (ii.). However, the nominal referring expression that is created is (iii.) a lexeme whose meaning is no longer an open elliptical formula, and whose category is fixed at the moment of coining. Thus, this function is used to coin new lexemes (in Svan) whose meanings are concretized from the potential range of the above open formula. This function is pure lexical derivation, and a few examples should suffice, though it appears to be productive. It would be bizaare and perverse to claim that our anomalous genitives are examples of pure lexical derivation, yet they share the same formal properties. Note that all such derived nominals (as well as all items not assigned a category lexically, such as adverbs) are uniformly of declension 6, which is the most productive declension in Svan:

kor (N) GEN *koräš* 'House' → *koräš* (N) 'Member of a family'
cxeḱ (N) GEN *cxeḱiš* 'Forest' → *cxeḱiš* (A) 'Wild'
iwaladäḡ (ADV) GEN *iwaladäḡiš* 'Always' → *iwaladäḡiš* (A)
 'Permanent'

(3.) While the previous two categories are primarily lexical, one deriving *ad hoc* referring expressions ('elliptical') the other actual lexemes ('lexical'), the third category is determined uniquely by the syntax. It has the syntactic properties listed above ((i.) and (ii.)), that is, taking an input expression (including [+Kin] nominals in Svan),

and returning a generic nominal expression. However, it has no semantic moment whatsoever: the input expression and the output may be said to have the same meaning. This is so because the conditions for this third function ('rank-shifting') are created and uniquely determined by syntactic or morphological environment. A category changing operation is performed which is isosemantic, but serves to produce a generic NP from one which is morphologically defective or categorially barred from a certain distribution. Thus, a proper name is 'rank-shifted' to a generic N, lacking the morphosyntactic feature [+Proper]. Since the environment is uniquely determined syntactically, and the meanings of N_{input} and N_{output} are seemingly identical, we would not want to say that this operation encodes anything at all.

The three functions I have listed share a common category changing core. However, their semantic moment, whether (1.) elliptical-derivational, (2.) lexical-derivational or (3) nil (as in rank shifting), is essentially a by-product of the ends to which they are put. In essence, this genitive freely creates generic nominals from anything else, for whatever purpose one might want new nominals for (coining, elliptical reference, or rank shifting). Just as the Greek substantivizing article can be elliptical, semi-derivational, or simply required by the exigencies of syntax (as when an infinitive is required to bear an oblique case), so too the genitive in Old Georgian or Svan may have each of these functions, and it appears to me that in Svan the role of the genitive in split locatives is to create a *generic Noun* from one whose categorial specifications ([+Proper], [+Kin]...) bar it from specific syntactic positions (e.g., [_{pp}-DAT P]). This generic N may now assume this position. However, insofar as this derivation is performed to meet the exigencies of grammar (precluding any possible contrast in this environment), and not at the behest of lexis, it is reasonable to assume that it has no effect on the meaning.

In this paper I have attempted to demonstrate that the phenomenon of 'Split Locativity' in Old Georgian and Svan is best handled in terms of a theory of morphosyntax (Silverstein 1976) that treats CASE-MARKING as a dependent variable computable from the interactions of a number of independent variables (including, but not limited to, propositional role and inherent denotational content of NPs). I have attempted to demonstrate that this theory allows us to make non-vacuous

predictions about the possible formal elaboration of such systems, in terms of both the number of bounded regions of the reference space that affect case-marking (unary, binary, ternary, to n -ary systems) and the sorts of formal mechanisms that tend to expound these bounded regions (CASE A, B, C). Further, it has been implicitly noted that there are no absolute a-prioristically definable boundary points for these regions, save that they must include certain features ([+Locative] for CASE B), whatever else they may spread to implicationally. I have tentatively proposed that the boundary for the CASE A and CASE C divide in Old Georgian is [\pm Human], while that for Svan is [\pm Kin]. Further, I have argued that the use of genitive with some adpositions in Old Georgian is rather different from its use in Svan: the former involves genitive and dative in paradigmatic contrast, while the latter hierarchically subordinates the genitive to the dative. Lastly, I have attempted to refine the earlier statements of Šanije (1976) and Abesadze (1984) on the pseudo-derivational use of the genitive, and have argued that insofar as the genitive derivation meets the exigencies of syntax rather than lexis or reference, and is uniquely determined by syntactic environment, the degenerative derivation in CASE C is isosemantic: it has only syntactic/categorical and not semantic, moment.

...

ACKNOWLEDGEMENTS

This is a substantially revised version of a paper delivered at the Conference on Non-Slavic languages of the USSR in Chicago (Spring 1991). I have benefited from discussion with the following people: Howard Aronson, Winfried Boeder, Michael Silverstein, Amy Dahlstrom, Lynn Nicholls, Kevin Tuite, Chris Roth and others. Responsibility for whatever errors I have made remain with myself. This paper is dedicated to the late Mair Morgan.

REFERENCES

- Abesadze, Nia. 1984. Adverbiale Bestimmungen mit Postpositionen und Postpositionalobjekte im Swanischen. In *Sprachen Kaukasiens*: 5-15 Jena: Friedrich-Schiller-Universität.

- Allen and Greenough. 1983 [1920]. *New Latin Grammar*. New York: Caratzas..
- Dixon, R.M.W. 1976. *A Grammar of Yidinj*. London: Cambridge.
- Goodwin, William Watson, and Charles Burton Gulick. 1981. *Greek Grammar*. New York: Caratzas.
- G&P. 1986. = Gudjedjiani, Chato and Mykolas L. Palmaitis. 1986. *Upper Svan: Grammar and Texts*. Kalbotyra 37(4). Vilnius: Mokslas.
- Haas, W. 1966. Linguistic Relevance. In: Bazell et al. (eds.) *In Memory of J.R. Firth*. London: Longmans:116-147.
- Halliday, M.A.K. 1961. Categories of the theory of grammar. *Word* 17:241-292.
- Hanks, William. 1990. *Referential Practice: language and lived space among the Maya*. Chicago: University of Chicago Press.
- K'ek'elije, K.' 1956. 'C'inaše,' 'tana' da 'zedā' tandebulta sint'aksuri punkciisatvis jvel kartulši.' In: *E'iudebi jveli kartuli lit' erat' uris ist' oriidan*:300-308. Tbilisi: St'alinis saxelobis Tbilisis saxelmc'ipo universit'et'is gamomcemloba.
- Larson, Richard K. 1985. Bare-NP adverbials. *Linguistic Inquiry* 16(4):595-621.
- Nichols, Johanna. 1983. On direct and oblique cases. *BLS* 9:170-192. Berkeley: Berkerley Linguistics Society.
- Šaniĵe, Ak'ak'i. 1976. *Jveli kartuli enis gramar' ik'a*. Tbilisi: Tbilisis universit'et'is gamomcemloba.
- _____, V. Topuria, and M. Guĵejiani. 1939. *Svanuri p'rozauli t'ekst'ebi, I. Balzemouri k'ilo*. Tbilisi: SSRK MASPG.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, pp 112-171 (Linguistic Series, 22.) Canberra, ACT: Australian Institute of Aboriginal Studies.
- _____. 1981. Case-marking and the nature of language. *Australian Journal of Linguistics*, 1:227-244.
- _____. 1984. Wasco-Wishram lexical derivational processes vs. word internal syntax. *Papers from the parasession on Lexical Semantics*:270-288. Chicago: Chicago Linguistic Society.
- _____. 1986. Noun phrase categorial markedness and syntactic parametricization. In: S. Choi, D. Devitt, W. Janis, T. McCoy, and Zh-sh. Zhang (Eds.), *Proceedings of the Second Eastern States Conference on Linguistics, The State University of New York at Buffalo, Buffalo New York, October 3-5, 1985*:337-361. Columbus: Ohio: State Univ. Department of Linguistics.
- _____. 1987. Cognitive implications of a referential hierarchy. In *Social and functional approaches to language and thought*. Orlando: Academic press: 125-164
- Smyth, Herbert Weir. 1976 [1920]. *Greek Grammar*. Cambridge, Mass.: Harvard University Press.
- Vogt, Hans. 1988 [1947]. Le système des case en géorgien ancien. (NTS XIV 1947:38-82) IN Even Hovdhaugen and Fridrik Thordarson (eds.) *Studia*

Caucasologia II: Linguistique caucasienne et arménienne. Oslo: Norwegian University Press.

RELATIVIZATION IN CHUKCHEE

Maria Polinsky

University of Southern California, Los Angeles

1. Introduction.¹

1.1. Goals. The goal of this paper is to analyze Relativization in Chukchee and then to propose a set of generalized rules for it. The Chukchee data presented below are from the author's own field notes and from the existing descriptions of Chukchee by Peter Skorik, Vladimir Nedjalkov, and Bernard Comrie (in case of disagreement between the published data and the field notes, preference was given to the latter).

1.2. General. Chukchee is a Paleo-Siberian language of the Far East, closely related to Koryak and Alyutor and also to Itelmen. Though the 1979 census establishes a population of 14,000 for Chukchee, most of these people speak Russian as their first language and have a passive or semi-active knowledge of Chukchee. According to my own informal estimate, the number of people who can really speak Chukchee is now less than a thousand, and language attrition rate is very high (see Polinskaja 1988 for a brief description of semilingualism in Chukchee).

Chukchee is a predominantly agglutinating, morphologically ergative language with the word order SOV/SVO (Agent Patient Verb/Agent Verb Patient). Chukchee has no articles and no grammatical gender distinctions. In the classification of Chukchee cases, it is relevant to distinguish between those cases that can encode terms ("term cases" below) and those cases that can only encode adjuncts ("non-term cases" below).

¹This paper is an extended version of the paper presented at the Chicago Conference on Non-Slavic Languages of the USSR in May 1991. The data were collected in the Soviet Union in 1985-1988. I am especially indebted to Larissa Kutgheut and Vladimir Raxtilin for their insights on the examples from their language. The interpretations are solely my responsibility.

The following abbreviations are used: ABS—Absolutive, AOR—Aorist, AP—Antipassive, DAT—Dative, DIM—Diminutive, ERG—Ergative, EV—epenthetic vowel, FUT—Future, INC—incorporated, INF—Infinitive, INTENS—Intensifier, LOC—Locative, NEG—Negation, NMLZ—Nominalizer, PART—Participle, POSS—Possessive; : indicates agreement; . separates the meanings in a portmanteau morpheme.

(1) Chukchee Case System:

| | | |
|-------------------------------|---|----------------|
| ERGative (-e ₁) | } | |
| ABSolute (-Ø, -ən) | | |
| DATive (-etə) | } | term cases |
| LOCative (-ək) | } | |
| ADiTive (-gʁit) | } | |
| ABLative (-gəpə) | | |
| MEDiative (-u) | } | non-term cases |
| COMitative (-e ₂) | | |
| SOCiative-Comitative (-ma) | } | |

Transitive verbs agree with the Subject and the Direct Object, and the subject agreement marker always precedes the object agreement marker, although there are a number of portmanteau agreement morphemes as well. Intransitive or detransitivized verbs agree with the subject only.

1.3. Framework. A brief mention of the syntactic framework is appropriate here. The approach used in this paper combines the theory of Functional Syntax (FS) with Relational Grammar (RG). Basic principles of FS are as follows:

1. Clause structures are analyzed in terms of mappings between semantic notions such as Agent, Patient, Goal, Beneficiary, Locative, Instrument and grammatical relations such as Intransitive Subject, Transitive Subject, Direct Object, Indirect Object, Oblique Object.

2. Each grammatical relation is characterized by two sets of properties:

- a specific form/linear position assigned to it
- form/linear position of other elements in the clause or across clauses that is determined by this grammatical relation.

3. In addition to the analysis of its grammatical structure, each clause/sentence type is analyzed in terms of its function, which includes:

- grammatical function (e.g., disambiguation)
- semantic function (e.g., expressing or leaving unspecified some propositional meaning)
- communicative function which is comprised of linking grammatical relations and such entities as Topic, Focus,

Contrastive, Situational.

One of the major differences setting FS apart from other frameworks is its claim that linguistic facts, in particular linguistic variation, have notional (semantic and/or communicative) motivation.

The analysis below also relies on two essential principles of Relational Grammar (RG), namely, multistratal representations and the notions Absolutive and Ergative.

While FS recognizes implicitly multiple levels of grammatical relations, RG has a much better developed set of principles according to which a nominal can bear distinct grammatical relations to a clause at different levels; this results in the recognition of such notions as "initial" and "final" grammatical relation.² For instance, an element can be initial Direct Object and final Subject, which is characteristic of the Passive.

The analysis assumed here also recognizes the Absolutive as the relation subsuming the Intransitive Subject and the Direct Object, and the Ergative as the relation representing the Transitive Subject, which is different from the above two. Thus, Absolutive and Ergative are treated not only as cases but as grammatical relations.

Section 2 discusses grammatical relations in Chukchee; sections 3 and 4 outline the conditions of term Relativization, section 5 outlines the conditions of non-term Relativization, and section 6 is a discussion of both Relativization rules.

2. Grammatical Relations in Chukchee. To adequately describe Relativization, it is necessary to examine the difference between Subjects and Direct Objects in Chukchee. In fact, the syntactic properties of Subject and Direct Object are largely similar, and on the whole Chukchee should be characterized as a language with prevalent syntactic neutrality, i.e., similar syntactic behavior of Subject and Direct Object. Both terms control null copies across conjoined clauses and under Equi; they do not differ in Reflexivization properties, as they both control the reflexive pronoun. Some arguments for the subjecthood of the Agent in transitive clauses come from Agreement, but since Chukchee Agreement also involves the opposition between direct and inverse forms, it will not be considered here.

Evidence for the subjecthood of the Agent in transitive clauses comes

²The similarities between the two approaches are outlined in Perlmutter 1981: 319-320.

from clause conjunction under *ənqorə* 'and, then.' Intransitive subjects in Chukchee control the null copy across clauses conjoined by *ənqorə*, cf. (2):

- (2) *yokwa-yŋ-ən yet-gʔi ənqorə Ø/ ŋinqey rəyegtelew-nin*
 loon-INTENS-ABS arrive-AOR.3SG and boy(ABS) save-
 AOR.3SG:3SG
 'The Great Loon came and saved the boy.'

The same holds for Transitive Subjects, while Direct Objects can only control a pronominal copy, and, in case of ambiguity, are coreferential to the full NP; cf. (3b, c):

- (3) a. *yokwa-yŋ-e ŋinqey rəyegtelew-nin*
 loon-INTENS-ERG boy(ABS) save-AOR.3SG:3SG
 'The Great Loon saved the boy.'
- b. *yokwa-yŋ-e ŋinqey rəyegtelew-nin ənqorə ekwet-gʔi*
 loon-INTENS-ERG boy(ABS) save-AOR.3SG:3SG Ø/ and
 Ø [= Loon/*boy] leave-AOR.3SG
 'The Great Loon saved the boy and Ø left.'
- c. *yokwa-yŋ-e ŋinqey rəyegtelew-nin ənqorə *Ø//?ətlon/ŋinqey*
ekwet-gʔi
 loon-INTENS-ERG boy(ABS) save-AOR.3SG:3SG and
 *Ø//?3SG/boy(ABS) leave-AOR.3SG
 'The Great Loon saved the boy and the boy left.'³

In (3b) the null copy is coindexed with 'the Great Loon' only, and in (3c) the coreference with 'the boy' is only possible if the full NP is repeated.

To recapitulate the treatment of case marking, the Intransitive Subject and the Direct Object are coded by the Absolutive, while the Transitive Subject is coded by the Ergative.

Grammatical relations other than the Subject and the Direct Object

³The use of the pronominal copy here is implausible, as Chukchee does not distinguish grammatical gender or human/non-human; therefore, *ətlon* can be coindexed both with 'the Loon' and 'the boy.'

cannot control null copies across clauses; this will be relevant for the discussion below.

3. Relativization of Subject and Direct Object. In earlier works on Chukchee, the formal aspect of Chukchee Relativization was briefly outlined in Comrie (1979) and Polinsky and Nedjalkov (1987: 242-244). Relative clauses in Chukchee invariably have the participle as predicate. Relativization strategies differ in the Declarative and Negative, which, for this reason, are analyzed separately.

3.1. Declarative Relative clauses. Intransitive Subjects relativize on the **-lʔ-** suffixed participle; this participle is usually described as the standard Active Participle (4b). Cf.:

- (4) a. *ŋinɣey pəkɪr-gʔi*
 boy(ABS) arrive-AOR.3SG.S
 ‘The boy arrived.’
 b. *pəkərə-lʔ-ən ŋinɣey*
 arrive-PART-ABS boy(ABS)
 ‘the boy who arrived/is arriving’

Direct Objects relativize on the **-yɔ-** participle, usually identified as the Passive Participle, cf. (5b):

- (5) a. *yokwa-yŋ-e ŋinɣey rəyegtelew-nin*
 loon-INTENS-ERG boy(ABS) save-AOR.3SG:3SG
 ‘The Great Loon saved the boy.’ (same as (1a))
 b. *rəyagtalaw-yɔ (yokwa-yŋ-e) ŋinɣey*
 save-PART loon-INTENS-ERG boy(ABS)
 ‘the boy saved (by the Great Loon)’

(5b) shows that Relativization of the Direct Object is possible with the Transitive Subject retaining the Ergative case; meanwhile, Relativization of the Transitive Subject by the same strategy as the Intransitive Subject is impossible; this is shown by (5c):

- (5) c. **rəyagtala-lʔ-ən (ŋinɣey) yokwa-yŋ-e*
 save-ACTIVE PART-ABS boy loon-ERG

'the loon that saved the boy'

As (5d) shows, it does not help if the Absolutive NP 'the boy' is left out:

- (5) d. **rəyagtala-lʔ-ən yokwa-yŋ-e*
 save-ACTIVE PART-ABS loon-ERG
 'the loon that saved the boy'

The only way to relativize the Transitive Subject is via Antipassive; cf. (5a) and (5e, f). In the Antipassive clause, the initial Direct Object becomes a non-term, and the Subject is coded by ABS. The Antipassive clause is illustrated by (5e) where the NP 'the loon' is in the Absolutive case and 'the boy' is in the Locative; for more on the Chukchee Antipassive see Kozinsky, Nedjalkov, and Polinsky 1988; and Polinsky and Nedjalkov 1987.

- (5) e. *yokwa-yŋ-ən ŋinqey-ək ine-nyegtele-gʔi*
 loon-INTENS-ABS boy-LOC AP-save-AOR.3SG.S
 'The Great Loon saved the boy.' (Antipassive)

- f. *yokwa-yŋ-ən (ŋinqey-ək) ine-nyegtelew-ə-lʔ-ən*
 loon-INTENS-ABS boy-LOC AP-save-EV-PART-ABS
 'the loon that saved the boy'

(5f) shows that the subject of the Antipassive clause relativizes on the -lʔ- participle, in the same manner as the Intransitive Subject.

Thus, Chukchee maintains a three-way distinction in Declarative clauses, namely between Intransitive Subjects that relativize on the so-called Active Participle, Direct Objects that relativize on the Passive Participle, and Transitive Subjects that relativize on the Antipassive Participle.

The relative clause agrees with its head in case and number, cf. (4b) and (6), where both the participle and the NP are in the Absolutive and Dative respectively.

- (4) b. *pəkərə-lʔ-ən ŋinqey*
 arrive-PART-ABS boy(ABS)

‘the boy who arrived/is arriving’

and:

- (6) *pkərə-lʔ-etə ɲenq-etə*⁴
 arrive-PART-DAT boy-DAT
 ‘to the arriving boy’

3.2. Negative Relative clauses. In negative relative clauses, the accessibility to relativization is the same, i.e., only Subjects and Direct Objects can relativize. However, the formal distinction is between Intransitive Subjects and Direct Objects on the one hand, and Transitive Subjects on the other; in other words, between the relations Absolutive and Ergative. Negative participles can take only the *-lʔ-* suffix. Thus:

- (7) *lonə-pəkərə-lʔ-ən ɲinqey*
 NEG-arrive-ACTIVE PART-ABS boy(ABS)
 ‘the boy who did not arrive’ (cf. (4b))
- (8) *luŋə-rəyegtele-lʔ-ən (yokwa-yŋ-e) ɲinqey*
 NEG-save-PASSIVE PART loon-INTENS-ERG boy (ABS)
 ‘the boy who was not saved (by the Great Loon)’ (cf. (5b))

As in declarative clauses, Transitive Subject is relativized via Antipassive. In both (9a, b), the NP ‘the loon’ is modified by the Antipassive negative clause. In the Antipassive clause, the initial Direct Object can be coded differently; thus, in (9a) the NP ‘the boy’ is coded by an oblique case, and in (9b) it is incorporated.

- (9) a. *luŋ-ine-nyegtelew-ə-lʔ-ən (ɲinqey-ək) yokwa-yŋ-ən*
 NEG-AP-save-EV-PART-ABS boy-LOC loon- INTENS -
 ABS
 ‘the great loon that did not save the boy’
- b. *luŋə-ninəqə-nyegtelew-ə-lʔ-ən yokwa-yŋ-ən*

⁴The comparison between (4b) and (6) illustrates morphophonemic changes associated with case marking; the discussion of Chukchee morphophonemic rules is beyond the scope of this paper.

NEG-boy(INC)-save-EV-PART-ABS loon-INTENS- ABS
 'the great loon that did not save the boy'

As in the Declarative, negative relative participles agree with their head noun in case and number.

Thus, Chukchee maintains a two-way distinction in the Relativization of terms in the negative, namely between Intransitive Subjects and Direct Objects that relativize on the negative *-lʔ-* participle, and Transitive Subjects that relativize on the negative Antipassive Participle. The use of the *-lʔ-* participle for Relativization of Patients (Direct Objects) allows us to argue against its characterization as the Active Participle. It seems more appropriate to characterize it as the Absolutive Participle.

4. Relativization of other terms. Chukchee only allows for Relativization of main or nuclear terms, i.e., Subject and Direct Object. In order to be relativized, the Indirect Object coded by the Dative case must advance to Direct Object, by means of 3-2-Advancement (Dative Shift in a different terminology); this advancement is limited to transitive clauses only, as the DO position has to be available for it. The Oblique Object coded by the Locative case⁵ advances to Direct Object only, through Oblique-2 Advancement. Cf. (10a,b,c) where (10a) illustrates a regular bitransitive clause, and (10b) illustrates the 3-2 Advancement:

- (10) a. *tumg-e keyŋ-ən akka-gtə təm-nen*
 friend-ERG bear-ABS son-DAT kill-AOR.3SG:3SG
 'The friend killed the bear for his son.'

In (10a), the verb agrees with the NPs in the Ergative and the Absolutive, 'friend' and 'bear' respectively. In (10b), the initial Direct Object is incorporated; the initial Indirect Object advances to Direct Object and now determines agreement⁶.

- (10) b. *tumg-e ekək kayŋə-nmə-nen*

⁵Nominals coded by oblique cases other than Locative, are not eligible for Oblique Advancement. See (1) above for the distinction between term and non-term cases.

⁶In this example, the form of the agreement marker does not change, as it still reflects 3 person acting upon 3 person; cf., however, (11a) and (11b) below.

friend-ERG son(ABS) bear(INC)-kill-AOR.3SG:3SG
 'The friend killed the bear for his son.'

The advanced Direct Object (*ekək*) can now relativize on the Passive participle:

- (10) c. *kayŋə-nmə-yə* (*tumg-e*) *ekək*
 bear(INC)-kill-PASSIVE PART friend-ERG son(ABS)
 'the son for whom the bear was killed (by [my] friend)'

In clauses such as (10b), the demoted (initial) Direct Object can no longer relativize.

In intransitive clauses, the initial Indirect Object coded by the Dative case may advance to Subject by means of Goal Advancement (see also Polinsky 1990). In (11a), the Goal is Indirect Object:

- (11) a. *ŋinq-eg-ti ətləg-etə et-ə-lʔet-gʔe-t*
 child-DIM-PL.ABS father-DAT INTENS-EV-come- AOR .3
 SG.S-PL
 'The father has many children (lit.: Many children came to the father).'

(11b) illustrates Goal Advancement: the initial Subject ('children') incorporates, and the Goal advances to Intransitive Subject; the respective NP ('father') is now coded by the Absolutive and determines verbal agreement, which is evident because of the loss of the plural marker on the verb.

- (11) b. *ətləg-ən ŋinqe-et-ə-lʔet-gʔe*
 father-ABS child(INC)-INTENS-EV-come-AOR.3SG.S
 'The father has many children.'

Now the advancee ('father') can relativize as a regular Intransitive Subject:

- (11) c. *ŋinq-et-ə-lʔe-lʔ-ən ətləg-ən*
 child(INC)-INTENS-EV-come-ACTIVE PART father- ABS
 'the father who has many children'

The difference in coding strategies notwithstanding, it has been so far demonstrated that Chukchee only relativizes Absolutives. This is true for the following initial Absolutives:

- Intransitive Subject
- Direct Object.

If the target of Relativization is not an initial Absolute, then the initial Absolute is demoted, and the target NP advances to the final Absolute. This happens with the following terms:

- Transitive Subjects that are initially Ergative and become final Absolutives by Antipassivization
- Indirect Objects in transitive clauses that become final Absolutives by 3-2 Advancement
- Oblique Objects in transitive clauses that become final Absolutives by Oblique-2 Advancement
- Initial Goals that become final Absolutives by Goal Advancement.

Thus, the following rule is proposed for Chukchee Relativization:

(12) In Chukchee, a final Absolute can be relativized.

It follows from (12) that only nuclear terms are directly relativized. Since Chukchee just does not allow non-terms to be promoted to Absolute, the restriction that this rule only apply to terms is unnecessary. To summarize, the relativized NP is always understood to function as the Absolute of the Relative clause.

E. Keenan (1985:155) indicates that "the question of which positions in a language can be relativized is not independent of the R[elative] C[laus]e forming strategy used." Indeed, Chukchee has quite a restrictive strategy of Relativization on participial clauses, as compared to finite verb relative clauses; the participles impose their inherent syntax on the relative clause. In addition, Chukchee has no relative pronoun or any other overt copy in the relative clause; the absence of such an element is always a constraint on relativizing certain positions in a clause.

Chukchee, however, has a separate rule for Relativization of adjuncts.

5. *Kin*-Relativization.

5.1. General description. In addition to the constructions shown above, Chukchee has a special strategy for relativizing adjuncts in

intransitive clauses.⁷ This strategy allows Relativization of adjuncts denoting location, time, attribute, or instrument of the given action or even a person co-participating in this action (Comitative). Such adjuncts are coded either by non-term cases shown above (1) or by the Absolutive + postposition, as in (13a). The predicate of the relative clause is the verb with the suffix *-kin-*; the clause also contains the initial subject coded as Possessor.⁸ It is important to note that the Relativization strategy for Indirect/Oblique Objects shown above does not work for adjuncts, since there is no independent advancement rule that could promote a non-term to a term.

- (13) a. *ʔorawetʔ-an iwtelet-gʔi (rʔet-yekwe)*
 man-ABS come down-AOR.3SG road(ABS)-along
 'The man came down (the road).'
- b.i. *iwtelet-kin ʔorawetʔ-en rʔet*
 come down-KIN man-POSS road(ABS)
 ii. *ʔorawetʔ-en iwtelet-kin rʔet*
 man-POSS come down-KIN road(ABS)
 'the road by which the man came down' (Relativization into PP)

Although both orders are possible, the order shown in (13bii) is preferable and more frequent, which is probably due to two related facts: first, this order keeps the two nominals apart, which facilitates processing; second, this order is consistent with the Intransitive Subject-Predicate order of finite intransitive clauses.

⁷The requirement that the clause be intransitive is relevant for matrix clauses only; it does not seem to hold for Relativization of adjuncts in embedded clauses. However, apparent cases of Relativization of adjuncts (such as presented in (i), where the instrumental NP 'shoulder blade' is relativized as downstairs Direct Object) can be explained with reference to the rules of Chukchee Causativization.

- (i) *ekke-te lʔu-nin ənpənač-en ketʔə-ken ʔəttəyotʔ-at parəʔlon-ən*
 son-ERG find-AOR.3SG:3SG old man-POSS remember-NMLZ
 ancestor-ABS.PL shoulder blade-ABS
 'The son found the shoulder blade which was to be used to make the
 old man remember his ancestors.' (Nedjalkov 1979: 249).

⁸ For a short description and examples of *kin* forms see Skorik 1977: 175-177.

Unlike the participial relative forms, the *kin*-form does not agree with the head noun in case (14a). The suffix *kin*, therefore, can be characterized as a nominalizer. It should be noted, however, that in spoken Chukchee the head noun tends to incorporate the relative clause if this noun is in any case other than the Absolutive. Cf. (14b), where the head noun is italicized:

- (14) a. ʔorawetlʔ-en iwtelet-kin rʔet-ək
 man-POSS come down-KIN road-LOC
- b. ʔorawtalʔa-wetletə-kin-e-rʔet-ək
 man(INC)-go down(INC)-KIN-EV-*road*-LOC
 'at/on the road by which the man came down

Terms cannot relativize on the *kin*-nominalized clause; this strategy is confined exclusively to non-terms. Cross-linguistically, nominalization seems to be a relevant strategy for extraction of adjuncts; structures, similar to (14b), are found, for instance, in Salish languages (Hukari 1977; 1979; Gerds 1988: 70-86).

5.2. The syntax of Adjunct Relativization. Chukchee has several nominalized constructions the analysis of which is beyond the scope of this paper; it should only be noted that while some Chukchee nominalizations can function as a complete sentence, the *kin* Nominalization cannot.

5.2.1. The grammatical relation assumed by the adjunct. For the *kin* Nominalization, the following analysis is proposed. The verb marked by *kin* becomes predicate of the nominalized construction. The nominal that corresponds to the subject of the finite clause is coded as Possessor (cf. above the Absolutive nominal 'man' changing into the Possessor nominal 'man's'). The initial adjunct acquires term status in the nominalized clause, and, after it becomes a term, is eligible for extraction, Relativization in particular.

The end result of the nominalization is quite similar to nominalizations in English, cf. (15b), though with the exactly opposite word order. The relative clause is nominalized, there is no copy of the head NP, and the subject of this clause is coded as Possessor.

- (15) a. we were discontent in winter

b. the winter of our discontent

The important question is what grammatical relations obtain in the clause that has the verb nominalized with *kin* as predicate. If we assume that *kin* Relativization is sensitive to the fact that only nuclear terms relativize in Chukchee (see (12) above), then following analysis could be proposed:

(16) In the nominalized clause, the adjunct becomes Subject

In other words, (13a) above may change into the unattested clause (13') which then yields (13b):

(13) a. ʔorawetlʔ-an iwtelet-gʔi (rʔet-yekwe)
man-ABS come down-AOR.3SG road(ABS)-along
'The man came down (the road).'

(13') *rʔet ʔorawetlʔ-en iwteletkin-Ø
road man-POSS going-PRED.3SG
'The road was man's coming down.'

(13) b. ii. ʔorawetlʔ-en iwtelet-kin rʔet
man-POSS come down-KIN road(ABS)
'the road by which the man came down'

5.2.2. Interaction with the initial and final Direct Object.

Advancement to Subject cannot occur over the existing Direct Object. In other words, only adjuncts in Intransitive clauses or detransitivized clauses are accessible to *kin* Relativization; i.e., the relative (nominalized) clause with *kin* has no Direct Object⁹. If the clause contains a direct object nominal, this nominal could either become Subject through the Passive, or become a non-term, through the

⁹This cannot be explained by a tentative general rule requiring that all nominalized clauses in Chukchee be intransitive, since nominalizations other than *kin* allow transitives, cf.:

(i) ətləg-ən resqik-wʔi aŋa-nwə ekək
father-ABS enter-AOR.3SG praise-NMLZ[PURPOSE] son(ABS)
'The father went in to praise his son.'

Antipassive. As was already mentioned, Chukchee has no Passive. Therefore, the only way "to get rid of" the initial Direct Object is to apply Antipassive prior to the *kin* Nominalization.

The examples above illustrated the intransitive clause; (17c) shows that the Relativization is preceded by the Antipassive (17b), with Direct Object demoted to Indirect Object in the Dative; (17a) shows that *winretək* 'help' is transitive.

- (17) a. *gə-nan enaalʔ-ən Ø-ra-winre-ŋən*
 2SG-ERG neighbor-ABS 2SG.FUT-FUT-help-3SG.FUT
 'You help the neighbor.' (Transitive)
- b. *gət enaalʔ-etə re-winret-gʔi*
 2SG(ABS) neighbor-DAT FUT-help-2SG.FUT
 'You help the neighbor.' (Antipassive)
- c. *gən-en winret-kin enaalʔ-etə wəlpə əŋŋe a-ntəyat-ka*
 2SG-POSS help-NMLZ neighbor-DAT spade(ABS) NEG
 NEG-leave behind-NEG
 'Don't forget the spade with which you will be helping the neighbor.' (Antipassive clause, Relativization into Sociative)

In (18b-c), Antipassivization results in the incorporation of the NP 'fish' into the verb.

- (18) a. *ərgənan ənn-ət na-ŋəttə-rkən-Ø-at*
 they(ERG) fish-ABS.PL INVERSE-look-PRES-3-PL
 'They are looking for fish.' (Transitive)
- b. *ətɾi ənnə-ŋəttə-rkə-Ø*
 -t
 they(ABS) fish(INC)-look-PRES-3-PL
 'they are fishing.' (Antipassive)
- c. *ənkə ənnə-ŋəttə-kin gətg-ən*
 there fish(INC)-catch-NMLZ lake-ABS
 'Over there is the lake where they fish.' (Relativization into PP)

The fact that Direct Objects are demoted in the *kin* Relative clause is easily explained by the impossibility for the adjunct to advance to Subject bypassing Direct Object. However, there is another, more interesting fact concerning Adjunct Relativization, namely, that some intransitive clauses do not allow *kin*-Relativization at all. Cf.:

- (19) a. *ʔorawetʔ-an tʔəl -gʔe aykol-ək*
 man-ABS be sick-AOR.3SG bed-LOC
 ‘The man lay sick in the bed.’
- b. **tʔələt-ken ʔorawetʔ-en aykol*
 be sick-NMLZ man-POSS bed(ABS)
 ‘the bed in which the man lay sick’

Syntactically, the ban on clauses such as (19b) can be explained in terms of the Unergative/Unaccusative distinction.¹⁰ According to the Unaccusative Hypothesis, some “intransitive clauses have an initial Direct Object but no initial Subject” (Perlmutter and Postal 1983: 69). Thus, *intransitive* verbs in the lexicon can differ according to the initial grammatical relation of their nominal argument; this argument can be initial (and final) Subject, as in the verbs ‘go,’ ‘work’ (unergative verbs) or it can be an initial Direct Object that subsequently ascends to Subject (unaccusative verbs). Semantically, this rule projects into the opposition of agentive verbs denoting controllable action and non-agentive verbs denoting uncontrollable action.

Assuming that the nominal argument of the Unaccusative is the initial Direct Object rather than Subject proper, the strategy of *kin* Relativization becomes quite consistent in that it only allows Relativization of adjuncts out of the structures that have the initial Subject. The relevant syntactic constraint on *kin* Relativization, which rules out Adjunct Relativization in Unaccusative clauses, is as follows:

- (20) Adjuncts relativize if they modify the clause that has

¹⁰For the Unaccusative/Unergative distinction in Chukchee and its independent semantic justification, see Polinsky 1990. For the general discussion of Unaccusatives/Unergatives, see Perlmutter 1978; see also Burzio 1986 for a similar analysis, in a different theoretical framework.

an initial Subject.

5.2.3. The status of the initial Subject. The next question in the analysis of Adjunct Relativization is the final status of the Possessor nominal which was the initial Subject. If (16) is correct, the assumption is that the initial Subject is demoted to a non-term. Indeed, it can be easily omitted: in quite a few cases, the *kin* clause has no overt Subject changed into Possessor. This occurs in those cases where:

(i) the Subject is generic, as in (21),

(ii) the Subject is one of the speech act participants, that is, first or second person, unambiguously recovered from the context, as in (22) and (23).

- (21) *teykew-kin ričit təlpʔi-gʔi*
 fight-NMLZ girdle tear-3SG.AOR
 'The fighting girdle (the girdle with which to fight) got torn.'
 (Relativization into Sociative)

- (22) *wayəŋqen welerkəle-kin rʔet*
 DEICTIC proceed-NMLZ road
 'Here is the road by which (you) should proceed.'
 (Relativization into PP; the Possessor/Agent is the addressee command)

- (23) *gəyiŋqewet-kin tumgətum qənwer tə-nŋə-gʔan yara-gtə*
 set stakes-NMLZ friend(ABS) finally 1SG-send-AOR.3
 house-DAT
 'I finally sent home the friend with whom I was setting stakes.' (Relativization into Comitative; the Possessor/Agent is the speaker)¹¹

However, the Possessor nominal retains some controlling properties; cf. (24) where 'man', though marked as Possessor, still controls Equi-NP-Deletion:

¹¹This example also shows that there are no animacy restrictions on the *kin*-Relativization. Though most adjuncts in the construction are inanimate nouns, animate nouns also occur, primarily as Comitatives.

- (24) iwtelet-kin ʔorawetlʔ-en ewne-kiməltet-ke rʔet
 come down-NMLZ man-POSS NEG-be late-INF road
 'the road by which the man came down in order not to be
 late'

Cross-linguistically, the retention of subject properties by the Possessor is not very surprising; cf. a similar, if marginally possible, behavior of the Possessor nominals in English nominalizations:

- (25) a. ʔ the time of his reading this document to please Joan
 was ill-chosen
 b. ʔ the methods of their advertising their products are annoying

It seems that this retention of controlling properties indicates that the Possessor is no longer Subject but remains Topic; however, this is an ad hoc suggestion which requires further study.

To summarize, the conditions on Adjunct Relativization in Chukchee can be formulated as follows:

- (26) a. only adjuncts are relativized through nominalization;
 b. adjuncts are relativized if they modify the clause that has the initial Subject (= (20));
 c. in the nominalized clause, the adjunct becomes Subject (= (16));
 d. in the nominalized clause, the initial Subject becomes Possessor.

5.3. The semantic motivation of the Adjunct Relativization rule.

In the Adjunct Relativization rule, the ban on Relativization out of unaccusative clauses seems unusual. The following is an attempt to argue for a semantic motivation of this ban and to propose a hypothesis that could be demonstrated or refuted based on cross-linguistic facts.

The explanation takes as the point of departure the hierarchy of semantic roles, arranged by their diminishing agentivity:

- (27) Agent > Patient > Instrument > Location > Comitative¹²

¹²The Addressee or Beneficiary that seems to belong with the Agent, rather than Patient, is excluded from the consideration for the sake of clarity.

The essential dichotomy is that of Agent and Patient, as these are two typical roles reserved for human participants. In the opposition Agent-Patient, the former has the ability to begin/continue/terminate a certain situation, in other words, is in control of the situation. Patient is defined as the participant whose state (dramatically) changes in the event described.

It is the participant that has the control over the situation that can create, use or manipulate other participants that are less actively involved in the situation. The agentive participant is relatively stable across time. Therefore, it can be used as the standard or reference point against which other participants are classified and interpreted. The stability of the Agent, therefore, ensures the unambiguous interpretation of other participants as well. Indeed, the prototypical narrative presents a single active participant or a limited number of active participants surrounded by many more inactive "props." The active participant, who is in control of the event, determines the localization and interpretation of the whole event.¹³

The objects that the active participant creates, destroys or modifies are less essential to the interpretation of the event and depend on the Agent. Therefore, the choice of Patient as the standard/reference point would lead to potential ambiguity: on the one hand, the Patient also ranks high on the agentivity scale (27) and can potentially control the participants coded as adjuncts; on the other, this participant is less time stable as he has no control over his own condition: the control is exercised by the Agent.

This reasoning allows us to conclude, if tentatively, that the hierarchy in (27) is relevant in terms of participant contingency: it allows us to predict what nominals in the clause could be used as reference points for the semantic and/or referential interpretation of other nominals.

There exists, therefore, a conceptual relationship between the agentive participant as the controller of the action and "user" of the adjunct

¹³It is important to emphasize the probabilistic nature of these observations; of course, such statements as:

- (i) The infinite Logos is inaccessible to man

are legitimate, especially in languages that have a long tradition of being used as a tool for philosophical deliberations.

referent, the action itself, and the item/participant denoted by the adjunct. In other words, the interpretation of the adjunct is contingent on the information about the central, agentive, participant in the situation. The knowledge of the agentive status of the respective participant rules out the potential interpretation that there is a third party controlling both this participant, as is typical with Patients, and the adjunct referent. This might explain why there is no contingency relationship between the Patient referent and the adjunct referent.

The relationship between the two referents is either permanent or situationally established. In the examples above, the function of the adjunct referent is unambiguously recovered, based on the lexical meaning of the verb; cf. a route in (13) and (22), an instrument of the traditional fight in (21), a coparticipant in the setting of stakes in (23), an instrument in (17c), a location in (18c). Importantly, the interpretation of the adjunct is "trivial": it is based on the encyclopedic knowledge of the ways certain objects are used. (Indeed, in theory a lot of things can be done at and with the road but the interpretation for the adjunct is always the most typical and predictable one.)

To summarize this section, the syntactic properties of adjunct extraction in Chukchee may have a cognitive explanation, related to the specific status of the Agent as the ultimate controller of the situation on the whole and of the adjunct referent in particular. Other situations where Agents are treated differently than other semantic roles are known from a number of languages (cf. Keenan 1984); however, the most common examples are those that distinguish transitive Agents, on the one hand, from intransitive Agents and Patients, on the other. Further cross-linguistic studies are necessary to find out whether the rules described above are idiosyncratic of Chukchee syntax or indicate some common tendency.

6. Term and Adjunct Relativization. At this stage, the question arises whether there is any rule that would unify (12) and (26), and if there is any reason for the fact that a language which is so restricted in its Relativization strategies allows for a "Relativization island" at the far end of the accessibility hierarchy that usually has the following form (adopted from Keenan and Comrie 1977):

- (28) Subject > Direct Object > Indirect Object > Oblique
Object > adjunct

As to the first issue, it seems that both (12) and (26) can be combined as complimentary elements of the Chukchee Relativization conditions, namely:

- (29) a. only nuclear terms are directly relativized;
 - a'. non-nuclear terms must advance to nuclear terms and then directly relativize;
 - b. only adjuncts are relativized through nominalization.¹⁴

The formulation of (29) resolves the issue of "Relativization islands" as (29 a') shows that no discontinuity occurs as regards the hierarchy of grammatical relations.

Combining the specific conditions outlined above (cf. (12) and (26)) into one larger rule would lead to a loss of information, rather than a concise description. Indeed, the strategies used to relativize terms and adjuncts differ dramatically, the former being participial relativization, the latter a nominalization with Subject changing into Possessor. While it is plausible to keep the two specific rules apart, it is also important that they both indicate the relevance of the relations Absolutive and Subject for the syntax of Chukchee.

7. Conclusion. This paper analyzes two separate Relativization strategies, namely, Term and Adjunct Relativization.

In Chukchee, only nuclear terms (Subject and Direct Object) relativize directly. A unified rule was proposed for Term Relativization, according to which the relativized NP should be the final Absolutive of the Relative clause. The rules that allow Indirect and Oblique Objects advance to nuclear terms, therefore, precede the Relativization of the respective NPs.

Adjuncts or non-terms relativize through nominalization which involves the coding of the initial Subject as Possessor. Adjuncts relativize only out of clauses with the initial Subject, which include unergatives and antipassives. A semantic explanation for this selective rule was proposed.

Though the two relativization strategies have different syntax and do

¹⁴These conditions are the same as obtained in Halkomelem for extraction which involves there Relativization, Clefting, Pseudo-Clefting, and Wh-Questions (Gerds 1988: 73, 59).

not seem to be governed by one rule, they are both sensitive to the grammatical relations Absolutive and Subject. With regard to Term Relativization, the latter feature is relevant in the choice of the participle; under Adjunct Relativization, only agentive Absolutive can be coded as Possessor in the Relative clause.

REFERENCES

- Burzio L. 1986. Italian Syntax: A Government-Binding Approach. Dordrecht: Foris.
- Comrie B. 1979. Degrees of ergativity: Some Chukchee evidence. In F. Plank, ed. *Ergativity*. London: Academic Press, 219-240.
- Gerdts D. 1988. Object and Absolutive in Halkomelem Salish. New York: Garland.
- Hukari T. 1977. A comparison of attributive clause constructions in two Coast Salish languages. *Glossa* 11, 48-73.
- Hukari T. 1979. Oblique Objects in Halkomelem. Papers from the 14th International Conference on Salish Languages. Bellingham.
- Keenan E. 1984. Semantic correlates of the Ergative/Absolutive distinction. *Linguistics* 22, 212-156.
- _____. 1985. Relative clauses. In T. Shopen, ed. *Language Typology and Syntactic Description*. V. II. Complex constructions. Cambridge: Cambridge University Press, 141-170.
- _____. and B. Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63-99.
- Kozinsky I., V. Nedjalkov, M. Polinsky. 1988. Antipassive in Chukchee. In M. Shibatani, ed. *Passive and Voice*. Amsterdam: John Benjamins, 651-706.
- Perlmutter D. 1978. Impersonal Passives and the Unaccusative Hypothesis. Berkeley Linguistic Society. Fourth Annual Meeting. Proceedings, 157-189.
- _____. 1981. Functional Grammar and Relational Grammar. Points of Convergence and Divergence. In T. Hoekstra et al., eds. *Perspectives on Functional Grammar*. Dordrecht: Foris, 319-352.
- _____. and P. Postal. 1983. The Relational Succession Law. In D. Perlmutter, ed. *Studies in Relational Grammar I*. Chicago: Univ. of Chicago Press, 30-80.
- Polinskaja M. 1988. Polujazyčie. In V. Belikov, ed. *Vozniknovenie i funkcionirovanie kontaktnyx jazykov*. Moscow: Nauka, 55-62.
- Polinsky M. 1990. Subject Incorporation: Evidence from Chukchee. In K. Dziwirek et al., eds. *Grammatical relations: A Cross-Theoretical Perspective*. Stanford: Stanford University Press, 249-264.
- _____. and V. Nedjalkov. 1987. Contrasting the Absolutive in Chukchee. In R.M.W. Dixon, ed. *Studies in Ergativity*. Amsterdam: North Holland, 238-269.
- Skorik P. Ja. 1961. Grammatika čukotskogo jazyka. I. Moscow-Leningrad: izd. AN SSSR.

_____. 1977. Grammatika čukotského jazyka. II. Leningrad: Nauka.

THE ASSIMILATION OF RUSSIAN VERBAL LEXICON IN MOKŠA MORDVINIAN

Stefan M. Pugh
Duke University

1.0. Until recently my research has dealt primarily with the subject of Russian-Balto-Finnic contact and interference, focussing specifically on the processes governing the assimilation of Russian verbs in such Balto-Finnic languages as Karelian and Vepsian, and to some extent Ingrian and Votic. The results of this research subsequently led to a similar study involving the use of Russian verbal lexemes in Komi; the similarities between the Balto-Finnic and Komi data were so striking that it seemed only natural to go even farther afield in Finno-Ugric, viz., to Mordvinian. The phonological and especially morphological peculiarities of Mokša and Erzä Mordvinian (both of which, like Komi, but unlike Karelian and Vepsian, are literary languages) promised to make this study something of a challenge.

This study is based on Mokša, primarily Mokša *dialect* material (*Očerki* I-IV), but I include some data from Erzä dialects (*ibid.*) as well as from literary Mokša (Keresztes 1990). The heavy reliance on transcriptions of dialect data is needed if we are to be serious about finding out how a Russian form becomes part of the Mordvinian system: forms culled from dictionaries (or exclusively from the literary language) represent lexicon that has been *normalized*; it is a traditional question of 'prescriptive' vs. 'descriptive' linguistics.

2.0. The role and influence of Russian in Mokša, as in other Finno-Ugric languages, has been investigated before. Gy. Décsy points out that this influence is most noticeable ('am ältesten und stärksten') in Mordvinian and Cheremis (Mari) because contact between the two linguistic groups can be traced to the 9th century, with the use of Russian loanwords dating approximately from the 10-11th centuries (Décsy 1988:632); older loanwords are cited in Feoktistov 1966:196, 218.

In spite of the long contact between these languages, and of the linguist's awareness of it, the processes that govern the assimilation of Russian forms to the Mokša system have not received the attention that the subject deserves. As is usually the case in such instances, it is customary for linguists to be content with the notion that a Russian

form will occur with native (L1) inflectional elements (endings); little or no attention is paid to the problem of morpheme boundaries: *how* does such a form occur with native endings? Before getting into Mordvinian, however, it is important to recall the model that I have found to be valid for both Balto-Finnic and Komi. In essence it is the following: the new verb is always based on the Russian non-past stem (all Russian verbs are identified as being either *-Vj-* or **non-*Vj-*** stems, and the latter all have a stem-marker *-i-*); frequently a morphological **+*VERB*** marker is affixed to the stem, followed by the Balto-Finnic or Komi verbal desinences. The important points here are (1.) that the Russian stem is a reality to speakers of these languages (they are, after all, bilingual) and (2.) that, in spite of (1.), Russian verbal lexemes often have to be additionally marked in some way to be identified as specifically *verbal* members of the L1 system. It is my intention, with this study, to attempt to find out if these two points are valid for Mokša as well.

A first glance at the Mordvinian verbs in question ('borrowings' or, if it is premature to term them 'loanwords', then perhaps 'usings': forms that are 'used' by a given speaker but not necessarily by the speech community as a whole) immediately makes one think that the situation is completely different from that of Balto-Finnic and Komi. In the latter two the Russian stem is clearly present in the new Finno-Ugric verb (audible in speech, visible in transcribed dialect texts), while it does not appear to be there in Mordvinian. In fact, I can only cite a very few forms where I could point out what appears to be a marker of a Russian verbal stem-type; there are even a few forms where a Russian marker of the non-finite forms (e.g., the infinitive) occurs, contrary to what we expect. Without any further investigation, the data then suggest that Mordvinian verbs formed from Russian lexemes are not **non-past stem-based** but **root-based**, or, in some instances, based on the Russian **past or infinitive**, contrary to what I found to be the case in other Finno-Ugric languages (see, e.g., Pugh 1990). Those few examples in which Russian stem markers *do* appear to be present can be explained as a function of the phonetic similarity (or even identity) of a given Russian marker and certain Mordvinian derivational formants and/or endings; forms exhibiting a suffix present only in the Russian past/infinitive can be explained in the same way.

In fact, unlikely as one might think it is, appearances are *not* deceiving in this case: study of all the forms presented in this study reveals an extraordinary departure from assimilatory patterns noted in previous analyses. It appears to be (and most likely is) the case that the significance of the Russian stem markers, so clear in Balto-Finnic and Komi, is altogether lost in Mordvinian. The 'Mordvinianization' of Russian forms has gone much farther in Mokša because the markers are for all intents and purposes gone, and the forms are adapted to one of three existing native verbal stem-types: *-a*, *-ə*, and *-#*.¹ This systemic difference (that is, Mordvinian vs. Balto-Finnic and Komi) must be ascribed to the early contact between the Mordvinian and Russian peoples; the Russian forms have simply had more time in which to become part of the L1 system. The forms entering the system since 1917 have been adapted in the same way, because the structural mechanism for the 'conversion' has already been established.

In order best to compare the Mordvinian situation with that of Balto-Finnic and Komi, I divide the body of borrowings and 'usings' into the categories 'Vj' and 'Non-Vj'; these designations reflect the nature of the stem-marking element in Russian (e.g., *-aj-*, *-ej-*, *-uj-*, *-i-*). Strictly speaking, of course, there is no synchronic Mordvinian basis for such a division: after the Mordvinian morphophonological processes have affected the shape of a given form, in the majority of cases, these forms can be reanalyzed as belonging to different Mordvinian stem-types which are completely independent of the original Russian stem-type. This does represent a point of agreement between the Mokša and Balto-Finnic processes, because the new verbs in the latter can also be described in terms of Balto-Finnic verbal stem types (I am not well enough versed in Komi to make the same claim for that language); the difference is, of course, that the original stem-markers are still identifiable in Balto-Finnic, whereas they are not in Mordvinian.

2.1 Non-Vj Verbs. The following represent almost all of the non-Vj forms that I have identified (the ones that have been left out are few in number and require additional study); differences or variations in grammatical endings (especially among vowels), are a function of the

¹I would like to thank Professor Daniel Abondolo of the School of Slavonic and East European Studies, University of London, for his insightful and extremely helpful comments on many aspects of this paper.

particular dialect in which these forms are used; some are the same as literary Mokša, others are not, and occasionally we find 'Erzä' endings occurring in 'Mokša' dialect texts. Consonantal palatalization (represented by ' and ") and place of stress are indicated as they are in the texts. In the interest of saving space, forms are not identified as, e.g., '3sg I past non-objective/objective "him/me",' as this information can be gleaned from the glosses. The first set of examples (a.) consists entirely of past tense forms, infinitives, and imperatives, because it is here that we best see the Mokša stem: the consonantal desinences are affixed directly to the vocalic stem without loss of the Mokša stem marker **-a-** (**-ä-**).² Note that palatalized consonants preceding front vowel stem-marking elements in the Russian base form are **retained** in Mordvinian:

- (a.) abída-z'ə 'he offended him' < obid'-i-
 fát"ä-s 'she grabbed' < xvat'-i-
 fát"ä-z'ə 'he grabbed him'
 fat'ä-z'ə 'he grabbed him' (lit. Mokša)
 kvat'a-š 'he recollected suddenly' also < xvat'-i-
 lád"a-maz' 'they set me up (put me in a good situation)' < lad'-i-
 pár'ä-z'ə 'he steamed' < par'-i-
 pr'ibáva-j 'increased' < pribav'-i- (morphologically a present tense)

²The following representative non-past and I past (non-objective) paradigms of literary Erzä and Moksha, based on Keresztes 1990, are provided for sake of comparison with the examples cited throughout this study (in this table the Moksha stem marker **-a-** [bold face] in the I past paradigm represents *both stem markers -a- and schwa*; in the present tense **-a-** is the marker of tense, not the stem-marker):

| NON-PAST | | I PAST NON-OBJECTIVE | |
|----------------|-------------|----------------------|-------------|
| <u>Moksha</u> | <u>Erzä</u> | <u>Moksha</u> | <u>Erzä</u> |
| 1 -a-n | 1 -a-n | 1 -a-n' | 1 -i-n' |
| 2 -a-t | 2 -a-t | 2 -a-t' | 2 -i-t' |
| 3 -a-j/-i | 3 -i | 3 -a-s' | 3 -a-s' |
| 1 -a-tamə | 1 -a-tano | 1 -a-mə | 1 -i-n'ek |
| 2 -a-tadə | 2 -a-tado | 2 -a-d'ə | 2 -i-d'c |
| 3 -a-Jt'/-iJt' | 3 -it' | 3 -a-s't' | 3 -a-s't' |

pros'ta-u-l't' 'they were/had been forgiven' (II past) < pros't'-i-;
 pros't'a-v-it' 'they are forgiven' (reflexive suffix -v-; in

preceding -v- > [u])

prošt'a-š' 'he excused himself'

sləža-s't' 'they served' < služ-i-

iz'ən' slúža 'I did not serve'

sləžá-n' 'I served'

páluč-ams 'to receive' < poluč-i-

prood'-amks 'to accompany' < provod'-i-

rozud'-amks 'to part, separate' < razvod'-i- (dial. roz-)

zaod'-amks 'to take someone or something to a place, drop off' <
 zavod'-i-

fát"-a-mak 'grab me!'

fát-ä-k / fát"-ak 'grab him!'

sroj-adu/-ada 'build' (2pl imperative, non-objective) < stro(j)-i-

pár'-a-mak 'steam me!' (in a *banja/sauna*) < par'-i-

Noteworthy among the forms listed above are those reflecting a non-palatalized root-final consonant. *Abid-* and *pribav-*: the former can probably be explained as influenced by the substantive *obida* [ab'ída], while the latter merely reflects the fact that [v'] is not part of the Mokša phonological inventory; in fact, all root-final labials resist and lose palatalization in the transfer from the Russian to the Mokša system. In *fat-ä-k* the frontness of [t'] is reinterpreted as frontness in the following vowel: [t'a] > [tä]; there appears to be no good reason for the hardness of [t] in *pros'taul't'* (but note the softness of preceding [s']).

In the following examples we find that a desinential vocalic element (1st or 2nd sg/pl present tense marker -a- or the 3rd person sg/pl ending -i-) has caused the truncation of the preceding stem-marking vowel; special note should be made of all desinences beginning with -i- outside of the 3rd person, because these are, strictly speaking, Erzä endings, even though they occur in so-called 'Mokša' dialects. Note also that we cannot tell in these cases whether a form has been assimilated as an -a-stem or as a schwa-stem verb because of the truncation:

- (b.) gatón-i-Jt' lit. 'they (things) get/are ready' < gotov'-i-guvər'-i 'he/she says' < govor'-i-lad'-i-z'i 'he aimed ('set up')' (the weapon) < lad'-i-mán'-a-sa 'I shall deceive him' < man'-i-man'-i-z'i/-i-z'ə 'he deceived (him)' < man'-i-man'-i-mik 'you [sg] deceived me' napus't'-i 'he/she will run away' < napust'-i-poluč-i-z'i 'he received it' < poluč-i-služ-y-n' 'I served' < služ-i-trud"-a-n 'I work, labor' < trud'-i-

The form *služ-y-n'* belongs to this group, even though the pre-desinential vowel is graphically represented as *y* instead of as *i*: this is merely the transcription of the allophone of /i/ that occurs following the hard husher [ž], reflecting either Russian or Erzä pronunciation.

The following set of forms represents clear evidence of assimilation as new Mokša schwa-type verbs; the retention of schwa is assured by the occurrence of the derivational suffix *-nda-* (and is not, therefore, limited to forms of the past, infinitive, or imperative):

- (c.) dubə-ndaz'ə 'he beat, cudgeled him' < dub'-i- 'tan' kəs't'ə-ndaz'e 'he christened him' < krest'-i-k'əs't'ə-ndaftəz'ə 'he had him christened' (causative suffix *-ft-*) mir'ə-ndás't' 'they made peace/reconciled' < mir'-i-rán'ə-ndamaz' 'they wounded me (I was wounded)' < ran'-i-rán'ə-ndaf '(was) wounded' past passive participle strójə-dndamə 'to build' < stro(j)-i-strójə-ndas' 'he built' strójə-dndavs' 'was able to build' (suffix *-v-* 'manage to do') (af) v'erə-ndaš't' 'they don't believe' < ver'-i-(af) vérə-ndan '(sometimes) I don't believe' vod'ə-ndakš'n'əs' 'bring to, place' < vod'-i- (frequentative-continuous suffix *-kšn-*) šut'ə-ndamanc 'to joke' < šut-i-šut'ə-ndal' 'he used to joke'

šət'əndat 'you joke'

Three forms occur with the apparent stem-marker (or pre-suffixal vowel) *-i-*: *s'el'-kəndatama* 'we shall settle (take up residence)' < *sel'-i-*, *spakó-ındat* 'you will calm down' < (u-)spoko(j)-i-, and *val'-in'dzi* 'he (over)filled it' < *val'-i-*. These might be explained in a number of ways, but one possibility is that they are interpreted as #-type verbs (i.e., where there is no stem-final vowel), in which case *-i-* is an epenthetic or fill-vowel which breaks up the consonant clusters arising after suffixation (*-l'dnd-*, *-jnd-*, *-l'ndz-*). In the second example, it is likely the presence of a preceding [j] that raises or keeps the vowel raised as [i]: in verbs like *uspokoit'* and *stroit'* there is a phonetic [j] between /o/ and /i/ in Russian. The question then arises as to why the same does not happen in the verbs based on *stroit'*: at this point the only explanation I can venture is that the difference is simply dialectal: the form with *-i-* occurs in the dialect of Temjaševo, while all three with schwa are found in the southwestern Srednevadskij dialect.

The following forms can be grouped together with the schwa-type verbs, because pre-desinential *e* can be interpreted as a dialectal phonetic variant of schwa:

- (d.) *lad'e-l't'* 'they set up, arranged' (Staroe Temjaševo)
napus't'e-s't' 'they flew/dashed upon' < (na)pust'-i- (Starye Turdaki)
služe-s' 'he served' < služ-i- (Staroe Temjaševo)
soglas'e-s' 'he/she agreed' < soglas'-i- (Starye Turdaki)
spor'e-s't' 'they argued' < spor'-i- (Starye Turdaki)
s'v'it'e-l't' 'they consecrated' < svjat'-i- (Staroe Temjaševo)

2.11 Consonant mutation. Several Russian verbal stem-types are characterized by consonant mutations: consonants (dental-alveolar stops and spirants) are mutated either in the 1sg only or throughout the entire non-past paradigm. In Balto-Finnic and in Komi the process of adaptation follows a logical progression: if a mutation occurs only in the 1sg in Russian, then a form is assimilated with the unmutated consonant (that which occurs in the majority of non-past forms); if it is found throughout the paradigm, then it is the mutated consonant

that appears. As speakers of Mordvinian are also speakers of Russian, it would not be unreasonable to expect the same pattern to hold for the assimilation of such verbs in Mordvinian. Most verbs with a mutation in the 1sg *do* occur according to the expected pattern, but occasionally they do not.

Three forms are attested in which an unmutated consonant should occur in the new Mordvinian verb (based on Russian verbs with a mutation in the 1sg only):

- (e.) *boroč*"a-mo 'to fight, struggle' < *borot*'sja (Erzä dialect); new a-stem
kantúž-indatan 'was contused, shell-shocked' < *kontuz*'-i-; new #-stem
nar'aža-s' 'he was dressed, dressed up' < *nar'ad*'-i-; new a-stem

The first of these three forms requires some comment, because the standard Russian stem is *bor*'- (*bor*'us', *boreš*'sja, etc.). G. I. Ermuškin (I believe mistakenly) cites this form as an example of the palatalization of the combination -t'-s- occurring in the Russian infinitive *borot*'sja (so ts > č); this, however, assumes that Mordvinian assimilates the verb based on the Russian infinitive, which is not the case. In fact there is a Russian dialect form *borotit*'sja (same meaning as literary *borot*'sja), of which the 1sg is *boročus*': this is the basis of the new Mordvinian verb. The form is *still* noteworthy, however, because the Russian dialect form only has palatalization in the 1sg. Cf. these exceptional forms with examples of stems reflecting the unmutated consonant (already cited): *kəst*- 'christen', *lad*'- 'arrange', *napus*'t'- 'fly upon', *prood*'- 'lead', *pros*'t'- 'excuse oneself', etc. The form *prošt*'aš reflects not the mutation of the Russian 1sg (which would be *prošč*-,) but the strong palatalization of [s'] > [š] brought about by the following palatalized [t'].

Similarly, four forms (from two Mokša dialects) based on verbs with consonantal mutation throughout the Russian non-past occur in the texts without mutation; all contain the root *-kaz-*, and all are assimilated as Mokša a-stems:

- (f.) *br'ikaza*-st' 'they punished' < *prikaž*-, inf. *prikazat*'

nakaza-z' 'they punished him' < nakaž-, inf. nakazat'

pr'ikaza-s' 'he ordered (gave an order)'

zakaza-s' 'he ordered (put in an order)' < zakaž-, inf. zakazat'

It is probable that these forms have been influenced by the presence in Russian of the corresponding deverbal nouns, all of which preserve the unmutated consonant: *nakaz*, *prikaz*, *zakaz* (cf. *abidaz* 'ə' above, in which hard [d] could be explained in the same way). We should not assume that they are based on the substantives, however; cf. *pr'ikazəvas*', based on the corresponding derived imperfective verb *prikazyvat*' (see 2.21 below); it is unreasonable to expect the Mordvinian speakers using these forms not to be aware of the mutated forms, because the Russian dialect forms have the same mutation pattern as literary Russian. It is also not plausible to try to infer that these forms are based on the Russian infinitive. Cf. also the literary Komi forms *nakažitny*, *prikažitny*, *zakažitny*. No forms of *-kaž-* or other verbs of this type (in the Russian system known as 'a-type' verbs, e.g., *-piš-*) are attested in the set of verbs collected for this study. It is impossible, on the basis of so few forms, to do more than to point out that the phenomenon exists; future studies will, it is hoped, bring to light other verbs of this conjugational pattern to which we can compare these forms.

2.2 Vj Verbs. Russian verbs of the 'Vj' types (aj, ej, uj < ova) are clearly assimilated as such in Balto-Finnic and in Komi, but they too lose their stem markers once they enter the Mordvinian system.

2.21 Aj. The vast majority of verbs belonging to the Russian 'Aj' conjugation often appear to retain the vowel *-a-*, but this vowel is not part of the original Russian stem-marker *-aj-*: it is, as in the case of non-Vj verbs, either a marker of the present tense, or the Mokša stem-marking element. Below I cite clear a-stem forms first (past tense, infinitives, imperatives):

- (g.) (iz'n') bəluč'á-dndakš'n'e 'I did not (used) to receive' < poluč-aj-dəbivá-dndams 'to obtain' < dobiv-aj-duma-s't' 'they thought (up)'
 duma-s' 'he/she thought' (lit. Mokša)
 dəmá-ftəms 'to make, cause to think'

óbža-j-n'ə 'I offended her' < obīž-aj-
 pr'ikazəva-s' 'he ordered' < prikazyv-aj-
 pr'ista-ndakt'n'əs't' 'they bothered, were a bother' < pristaj-
 (pristavat')
 pr'ista-ndakš'əs' 'he was a bother'
 pərváža-maz' 'they accompanied me' < provožaj-
 rabota-ma/-ms/-mə, robota-mu, róbəta-ma 'to work' < rabot-aj-
 rabota-ftəl' 'he caused/made them work'
 róbəta-ftəms 'to force to work'
 róbəta-n' 'I worked'
 róbəta-kšən' 'I worked a lot'
 rabota-kšəs' 'she worked a lot'
 rabóta-me 'we worked'
 rəgá-ms 'to scold' < rog-aj-
 s'p'eva-ndas' 'they were singing' < -spev-aj-
 sroga-j-emks 'to plane, shave' (Erzä dialect) < strog-áj-
 stara-ndaz'e 'he tried (it)' (lit. Mokša)
 s't'r'eč'a-mə 'to meet' < vstreč-aj-
 s't'r'eča-š't' 'they met'
 s't'r'eč'á-ndaz' 'they were meeting him'
 ven'c'á-dndatam 'let's get married!' < venč-aj-
 ven'c'á-dndas't' 'they got married'
 ven'c'a-dndame 'we got married'

The following verb forms are either clearly part of the Mokša schwa-type conjugation (note especially those forms with the derivational suffix *-nda-*), or they have lost the stem marking element (*a* or schwa) to truncation because of a following desinential vowel:

- (h.) dúm-əndas't' 'they thought'
 dum-əndas't' 'they think' < dum-aj-
 dum-əndás' 'he/she thought'
 dum-i 'he/she thinks'
 dum-it' 'they think' (a 'Mokša' dialect, but here Erzä end.)
 dəm-át 'you think'
 dym-an 'I think'

dum-əndáj 'he/she thinks'
 gul'-adndaj't' 'they stroll' < gul'-aj-
 ot'v'ič'-ej / at'v'ičj-ej / -aj 'he/she answers' < otveč-aj-
 otv'eč-i 'he/she answers'
 óxədnd-aj 'says "oh!"' < ox-aj-
 plot'n'ič'-əndamə 'to work as a carpenter' < plotnič-aj-
 róbat-aj 'he/she works'
 róbat-aJt' 'they work'

Several of these examples require some comment. *Otv'eči* could be assigned to another category, viz. that of Non-Vj verbs in which an unexpected mutated consonant occurs in the Mordvinian form (cf. Russian *otvetit'*-*otveču*, *otvetiš'*), but the evidence of a parallel form such as *dumi*, which can only be based on the -aj- verb *dumaj-*, suggests that it ought to be included here as an Aj verb. Next, the example *pərvážamaz'* might be interpreted in the same way, viz. as an example of an unexpected Non-Vj-type mutation (*provodit'*-*provožu*-*provodiš'*). In fact, no matter which interpretation we choose, the place of stress causes problems: as a verb of the Vj type (derived imperfective *provóžát'*-*provóžaju*) the vowel preceding the husher might be perceived in Mordvinian as a schwa, and would not attract the stress; as an i-type (non-Vj) the verb has both stressed [o] and schwa. In my opinion this Mordvinian verb does reflect the derived imperfective, but, as is sometimes the case in that language, it is a southern Russian/Ukrainian feature that is reflected: the corresponding Ukrainian derived imperfective is *provádžudvaty*; in other words, there is now a model with stressed [á] on which the Mordvinian verb can be based. Finally, we see here several forms with one of the variant Mordvinian 3sg non-past endings, -a-j (which can occur under or outside of stress): *óxəndaj*, *róbataj*; striking about this is the phonetic identity of this morpheme with the Russian stem marker -aj-; nevertheless, it is the Mordvinian identity of this element that is primary, because there is no 3sg non-past desinence consisting of # (zero) that would occur in the verb without a Mordvinian stem-marking element. Only two of the forms appear to reflect the Russian stem marker: *óbžajn'ə* and *srogajemks*; the latter is an Erzä dialect form, in which the element -aj- might be retained (1) because it is stressed in Russian (as is -aj- in

Rus. *obižaj-*), and (2) because the sequence *-a- + -emks* naturally supports the presence of *-j-*.

2.2. *Ej* Verbs. There are three forms that are classified as *ej*-verbs in Russian, and all three are based on the same Russian verb, *žalej-* 'have pity, be sorry'. The first two examples have apparently undergone the same changes as many of the *Aj* verbs, viz. *-ej-* is lost and the verb is assimilated as a Mokša *a*-stem: (*u*)*žal' a-samaz'* 'they pity me.' The third, however, attested in an Erzä dialect, reflects the original Russian stem marker, albeit somewhat changed phonetically: *žal' ij-iže* 'he pitied'; it is possible that the original Russian stems are better preserved as a whole in that language (cf. *srogajemks* above), but without much more research such an assumption would remain mere conjecture.

2.3. *Uj* (*ova*) Verbs. One of the most striking groups of borrowings or usings is that of the 'ova' verbs, the non-past marker of which is *Uj* (*sovetovat' -sovet-uj-u* 'advise'): verbs of this type appear to be based on the Russian non-finite forms, that is, with the suffix *-ov(a)-*. This is contrary to all expectations, especially when considered in light of Balto-Finnic and Komi, in which the non-past suffix *Uj* is clearly present (Kar. *torguija* 'trade,' Komi *piruitny* 'feast'):

- (i.) *br'ézg-əv-ams* 'to be squeamish' < *brezg-uj-pir-ov-amu* 'to feast' < *pir-uj-rad-əv-aš't'* (*ezənza*) 'they are happy for him' < *rad-uj-rad-uw-aš* 'he was happy' (an Erzä dialect)
rad-u-as 'he was happy'
r'emon't'ir-uv-amks 'to repair' < *remontir-uj-(kak) s'l'ed-əv-ət* 'as it should be' < *sled-uj-talk-uv-amks* 'to explain, explicate' (an Erzä dialect) < *talk-uj-*

These forms are not, in my opinion, based on the Russian non-finite stem *-ova-*, rather they appear to present us with a process of formal interaction between the Russian and Mordvinian systems, in which both phonology and word-formation are involved.

(1) In phonological terms, we can see here an interaction between the glides *j* and *v*: a relationship between them exists, as it happens, in the source language as well as in the receiving language. In Russian the much written-about relationship and alternation is most striking in

the area of verbal derivation, e.g., in the derivation of new imperfectives, not to mention within the *-ova-* conjugation itself, in which there is an opposition (or alternation) *infinitive v : non-past j*. In Mordvinian, according to Paasonen 1893, *v* can become *j* after front vowels; the opposite development is also found, apparently triggered by the presence of preceding **back** vowels (Paasonen 1893:32-33). The presence of [u] before [j] in the Russian non-past could therefore provide the impetus for the change *j > v*; then [u] before [v] can either be retained, as in four of the examples, or lowered to [o] (as in *pirovamu*) and then unrounded to schwa, as in the remaining examples. We note in *raduwaš* the development of the labio-dental [v] into the bilabial glide [w], which can then be lost, as in *raduas'* (the presence of the former argues against positing a lost *j* in the latter). Regarding the form *sledəvət*, it appears to consist of a contamination of the Mordvinian *-v-* formation and the Russian form *sledujet* (note the final -t). I should add a parenthetical note here regarding the relative proximity of Mordvinian to the Ukrainian speech area. In Ukrainian the verb type corresponding to Russian *-ova-* is *-uva-*: *tolkuvaty*, *raduvaty*, *remontuvaty*; the similarity of these infinitives to some of the Mordvinian forms can only be interpreted as coincidental, however: the non-past stem of Ukrainian *-uva-* verbs is the same as the Russian (*-uj-u*, etc.). Cf. especially Mordvinian *remontiruv-*, which appears to indicate Russian-Ukrainian interference: R. *remontirovat'* (in which Rus. *-ir-* < German < French), but U. *remontuvaty*. The phonological explanation appears to be much stronger than the possible role of Ukrainian influence, although the existence of the Ukrainian forms could be cited as additional support for the changes.

(2) There is strong reason to believe that Mordvinian word-formation has also played a significant role in establishing Russian *-ova-* verbs in Mordvinian in the shape in which they are cited here. Specifically, I refer to the presence and common use of the multipurpose derivational suffix *-v-* in Mordvinian: this suffix can express reflexivity, passivity, the possibility that an action can take place, as well as the sense that an action is taking place against the will of the actor (Paasonen:1893:32). The fact alone that such a versatile word-formational element exists should not allow us to presume that Mordvinians interpret the [v] of *-ova-* verbs as **being** this suffix; its existence might, however, support

the initially **phonological** change of [j] > [v] by endowing the resulting sound with some **word-formational and semantic** *raison d'être*.

3.0. +VERB Markers. The last question to be addressed is the need, seen in Balto-Finnic and in Komi, for the presence of some Mordvinian element in assimilated Russian forms that identify them as +VERB (i.e., a marker of morphological class). Clearly, the Balto-Finnic/Komi solution, that is, to use a suffix expressing transitivity or intransitivity, is not necessary: the presence in Mordvinian of the objective and non-objective conjugations generally excludes this as an option. Of course, as we have seen, a suffix such as -v- can be present if a verb is strongly reflexive (cf. *pros't'avit'* in (a.)). Most of the verbs occur only with grammatical desinences, and therefore appear to be able to enter the system without special markers; nevertheless, judging by the large number of forms found *with* such suffixes (and often with more than one), the process of *Aktionsart* suffixation is clearly vital to the system, as it is in Balto-Finnic (cf. Pugh forthcoming). It is worth pointing out, however, that fully 36 of the verbs on our list, or about one third of the total, are attested with the Mokša frequentative suffix -nda-, or with -dnda-, apparently consisting of -d- (expressing the momentaneous or one-time nature of an action) + -nda-. In some of the instances, -nda- does indeed express frequency or length of an action, while in others it seems not to do so. We notice also that this suffix does **not** occur in the -ova- verbs cited here, presumably (if we follow this line of reasoning) because -v- is a clearly verbal element, so that the presence of -nda-—other than as an *Aktionsart* suffix—would be redundant. This question should be addressed further when there is a larger data base with which to work (involving the scanning of more texts, and perhaps an extension of the study to Erzä Mordvinian). Suffice it to say that the frequency with which this suffix occurs in borrowings and 'usings' from Russian is somewhat suspicious, and suggests that a +VERB marker is at times necessary, perhaps for newer loans.

4.0. Conclusion. To summarize: what *appeared* to be absent from the Mokša verbs (a Russian stem marker) *was* in fact absent. The process of morphological adaptation goes quite far in the Mordvinian (Mokša) system, especially in comparison with Balto-Finnic: Russian verbal lexemes are reinterpreted as belonging to one of the natively

occurring verbal stem types, in the process losing the stem-marking elements characteristic of the Russian non-past conjugation. A special place in the Mokša system is occupied by the one Russian stem-type which appears to retain some form of a Russian suffix when used in Mokša: verbs of the *-ova-* / *-uj-* type are subject to a Mokša phonological alternation (j-v) which, in turn, facilitates interaction with the word-formational system (i.e., identification of -v- with the Mordvinian derivational morpheme). The net result, of course, is that the new forms have lost the Russian non-past marker *-uj-*.

In comparison with our Balto-Finnic data, the Russian stem is less stable and significant in Mokša. The greater resistance of the Moksha system to **structural** influence from the Russian language, e.g., compared with Komi, a language as established as Mordvinian, can be ascribed to the greater length of time of contact with Russian. The loss of the 'Russian' identity of borrowings and usings has been assisted by the potential for the confusion of such Russian stem markers as *-aj-* and *-i-* with native Mordvinian inflectional morphemes, because of the phonetic identity of these elements (compare these Russian markers with the 3rd person singular morphemes in the table provided in footnote 2). Whatever the reasons for the changes to the Russian stems, however, the result remains clear: Russian lexemes undergo a greater degree of assimilation to the native Finno-Ugric system in Mokša than they do in Balto-Finnic and in Komi.

REFERENCES

- Décsy, Gy. 1988. "Slawischer Einfluss auf die uralischen Sprachen," in D. Sinor, *The Uralic Languages: Description, History, and Foreign Influences*. Leiden: Brill, 616-37.
- Feoktistov, A. P. 1966. 'Mordovskie jazyki,' in K. E. Majtinskaja, ed., *Finno-ugorskie i samodijskie jazyki—Jazyki narodov SSSR*. III. Moscow, 172-220.
- Keresztes, L. 1990. *Chrestomathia Morduinica*. Tankönyvkiadó, Budapest.
- Očerki. 1961-66. = *Očerki mordovskix dialektov*, I-IV. Mordovskoe knižnoe izdatel'stvo, Saransk.
- Paasonen, H. 1893. *Mordvinische Lautlehre*. Helsingfors.
- Pugh, S. 1990. "Die morphologische Assimilation russischer Verben im Ostseefinnischen und im Komi." *Congressus Septimus Internationalis Fenno-Ugristarum*, Debrecen.

- _____. (forthcoming). "Balto-Finnic Verbal Derivation and the Assimilation of Russian Lexemes." *Linguistica Uralica*, Tallinn.

THE ROLE OF THE ZERO ACCUSATIVE IN KOMI FOLK TALES

Carol Rounds
Columbia University

Komi is a language of the Permian branch of Finno-Ugric. The Komi language actually breaks down into two main language groups: Komi-Permiak and Komi-proper. It is the latter that we will be dealing with here.

Direct-object marking is a complex operation in several Finno-Ugric languages, but particularly so in Komi, for it boasts five separate accusative cases. The use of each of them will briefly be described; the body of the paper, however, will be restricted to the comparison of the distribution of only two of the accusative cases. Finally, there will be a discussion of word order in Komi in relation to direct-object marking.

Direct-Object Marking. The five direct object markers are the following: a zero-accusative or nominative, the simple accusative *-ēs* (where *ē* is a schwa), the personal pronoun accusative *-ē*, and two possessive accusatives, second person singular possessive *-tē*, and the third person singular possessive *-sē*.

The distribution of these five markers is somewhat problematic, but one iron-clad generalization can be made. The *-ē* accusative is found only on singular personal pronouns. The first and second persons singular employ *-ē* exclusively. The third person singular is marked by either the *-ē* or the simple accusative *-ēs*, the latter being the more common form. Plural personal pronouns also exhibit the simple accusative form *-ēs*.

Both the second and third person possessive accusative cases are used to mark definite direct objects, just as the second and third person possessive endings are used to mark any definite nominal. The distinction between direct objects and other functional categories in the noun is that, for the most part, the possessive suffixes are suffixed on the nominal and followed by another case, whereas in the direct object the possessive and the accusative are fused into one marker. The third person possessive accusative has a very generalized use; the second person possessive suffix, however, has a more discourse-oriented function. It is found primarily in dialogue, i.e., primarily when there a first and second person axis of discourse. In narrative, the third person

possessive accusative is found where the second person is found in dialogue. For example:

Dialogue:

- (1.) “no dida kosamesańjas vaje žë menym lapítitë!”
 well sheepstall landlady-pl-nom bring-imp interj sg1-dat sandal-
 sg2-acc¹
 “‘Well, sheepstall landlady, bring me *the* sandals!’”

Narrative:

- (2.) sijë vremjanas batë këtšsë kulis.
 this time father-sg1px rabbit-sg3px-acc skin-sg3-pret
 ‘During this time my father skinned *the* rabbit.’

The direct objects with possessive suffixes are definite: the second person possessive merely fulfills the discourse function of including the second person in the dialogue. The second person possessive suffix can also be found in narration when the narrator wants to include the listener in his story. For example:

- (3.) on ušpejit ješë ledžny vugyrtë samnas, kydži tentšyd kvaķitasny
 da kutasny nuny vugyrtë.
 neg-sg2 manage still let down-inf hook-sg2-acc bait-s3px-ill
 when sg2-dat grab-pl3 and begin-pl3 pull -inf hook-sg2-acc
 ‘No sooner do you manage to set *the* (‘your’) hook when they
 grab and pull *the* (‘your’) hook.’

The two remaining accusative markers, the zero-accusative and what is here referred to as the simple accusative must be dealt with together. Although the distribution of these two cases appears at first glance to be unsystematic, some generalizations can be drawn.

¹Abbreviations used are: *acc*—accusative, *all*—allative, *comit*—comitative, *dat*—dative, *ill*—illative, *imp*—imperative, *inf*—infinitive, *interj*—interjection, *nom*—nominative, *pcp*—participle, *postp*—postposition, *pret*—preterite, *px*—possessive suffix, *refl*—reflexive.

The simple accusative tends to be used for (a.) animate direct objects:

- (4.) a askinas dođđali vëlës.
 and morning sled-sg1-pret horse-acc
 'And in the morning I sledded up the horse.'

and (b.) plural direct objects:

- (5.) mijë bätëkëd zaradítimë pištšáljasës.
 pl2-nom father-sg1px-comit load-pl1-pret rifle-pl-acc
 'My father and I loaded the rifles.'

The zero-accusative tends to be used in the following four scenarios:

(a.) generic constructions:

- (6.) da sy bëryn jëšë unays-na me vetli tseri kyjny.
 and that after-postp still many times sg1-nom go-sg1-pret fish-
 0acc hunt-inf
 'And after that I still went fishing (fish-hunting) many times.'

(b.) for inanimate singular indefinite objects:

- (7.) vëtšis ašlys tšom.
 make-sg3-pret sg3refl shelter-0acc
 'He made himself a shelter.'

(c.) to express a partitive notion:

- (8.) tšukërtis koš pes.
 collect-sg3-pret dry wood-0acc
 'He collected (some) dry wood.'

and (d.) after numerals:

- (9.) loktigën mijë lyjimë kyk dozmër.
 go-pcp pl1-nom shoot-pl1-pret two woodgrouse-0acc

'On the way we shot two woodgrouse.'

The tendency for objects occurring after numerals to appear with the zero-accusative is not as strong as the tendency for human animates to appear with a simple accusative; thus human direct objects occurring after a numeral are marked with the simple accusative.

(10.) kujimly vėlėma šetėmaėš zadañė vajny kujim mortės ta-vun keža.

three-all be-sg3-pret give-pcp duty-nom bring-inf three man-acc
that day until-postp

'Three of them were given the duty to bring three men for that day.'

As was stated above, the zero-accusative tends to mark (or, perhaps more accurately, *does not* mark) indefinites. If we think in binary terms this would imply that the simple accusative marks definites. We have already seen, however, that the possessive accusatives fulfill the definite-marking function. Thus, this binary distinction for definite/indefinite does not work here—rather, a further distinction must be made. If definites are marked with the possessive suffixes, then "non-definites" are marked with the zero and simple accusatives. The next distinction to be made within "non-definites" is topical vs. non-topical expressions. The zero-accusative is reserved for inherently non-topical direct objects and the simple accusative for topical objects. Because the complements of verbs of generic constructions are also non-topical, direct objects of generic constructions receive zero-accusative marking. A scale of topicality is found in the breakdown of which case marking corresponds to which direct object. A scale of increasing topicality in noun classification may be represented as the following:

singular/indefinite/inanimate >> plural >> animate >> human

The higher on the scale of topicality, the more likely we are to find simple accusative marking. Furthermore, we may assume that elements high on the topicality scale tend to be definite as well and therefore often receive the definite markers of the possessive suffixes. Indeed,

there is quite an overlap in Komi, where plurals and animates select either the simple accusative or possessive suffixes. The actual distribution of possessive vs. simple accusative case marking remains elusive at this time. We can only refer to the overlap in the topicality and definiteness of the direct object in question.

Word-order. In tandem with case marking, the word order of sentences must also be discussed. Is there a distinction in case marking that corresponds to word order? The answer, again, lies in the topicality scale. Komi, as with the other non-Baltic-Finnic languages in Finno-Ugric, is usually classified as an SOV language. There is, however, a high incidence (approximately 50%) of SVO word order. Again, for the most part I will restrict my discussion to the word order of sentences with only zero- or simple-accusative case marking.

Generic constructions tend to an SOV word order. This can be seen in example (5.), repeated here:

- (5.) da sy bëryn jëšë unays-na me vetli tšeri kyjny.
 and that after-postp still many times sg1-nom go-sg1-pret fish-
 0acc hunt-inf
 'And after that I still went fishing (fish-hunting) many times.'

If, however, there is a one-time occurrence or performance of a generic activity the word order changes to SVO. There is no change in case marking —only word order. For example:

- (11.) a atšys kutis kuľny këtš.
 and sg3-refl-nom begin-sg3-pret skin-inf rabbit-0acc
 'And he himself began to skin (a) rabbit.'

The activity of the generic construction has here been individuated and has consequently lost a degree of "genericness." This slight change in individuation refers, however, not to the complement of the verb, but rather to the performance of the activity itself, thus the change in marking is found in word order and not on the verbal complement.

For word-order tendencies of direct objects marked with the simple accusative, we must refer back to the topicality scale. If a direct object marked with the simple accusative is high on the topicality scale, it

may occur in a pre-verbal position. If it is low on the topicality scale it, occurs post-verbally. Thus, personal pronouns must occur pre-verbally, animates marked with the simple accusative can occur either pre-verbally or post-verbally and plurals and inanimates marked with the simple accusative must occur post-verbally. For example:

(a.) personal pronouns (only OV word order):

- (12.) *i ponjasys sijēs ku'isny kurttsalny*
 and dog-pl-sg3px sg3-acc begin-pl3-pret bite-inf
 'And the dogs began to bite him.'

(b.) animates (with VO word order): (example (4.) from above):

- (13.) *a askinas dođ đali vėlēs.*
 and morning sled-sg1-pret horse-acc
 'And in the morning I sledded up the horse.'

(c.) animates (with OV word order):

- (14.) *barin petukēs kutas da ēšjas dinē jērtas.*
 master hen-acc capture-sg3 and ox-pl-nom to-postp lock-sg3
 'The master captures the hen and locks (it) up with the oxen.'

(d.) inanimates (only OV word order): (example (7.) from above):

- (15) *vētšis ašlys í šom.*
 make-sg3-pret sg3refl shelter-0acc
 'He made himself a shelter.'

We may conclude that the key to understanding the distribution of the five accusative cases resides in the topicality of the direct object itself. Furthermore, the word order also reflects the same scale of topicality.

REFERENCES

(SUST = *Suomalais-Ugrilaisen Seuran Toimituksia*. Helsinki.)

Baker, R. 1985. *The Development of the Komi Case System*. = SUST 189.

Givon, T. 1984. *Syntax: A Functional-Typological Introduction*. vol.1. Amsterdam: Benjamins.

_____. (1990) *Syntax: A Functional-Typological Introduction*. vol.2. Amsterdam: Benjamins.

Kokkonen, P., ed. 1990. *Syrjänische Texte*. [Collected by T. E. Uotila] = SUST 202.

Rédei, K. 1978. *Zyrian Folklore Texts*. Budapest: Akadémiai Kiadó.

Schlachter, W. 1960. *Studien zum Possessivsuffix des Syrjänischen*. = *Finnisch-Ugrische Studien* 3. Berlin: Akademie-Verlag.

Wichmann, Y. J. (1916) *Syrjänische Volksdichtung*. = SUST 38.

ON THE LEXICALIZATION OF CLASSICAL ARMENIAN VOWEL EPENTHESIS¹

Frederick W. Schwink
University of Texas at Austin

0.0 Introduction. Epenthesis is a process by which a segment is added to a pre-existing phonological sequence without etymological justification. One may think, for example, of the English *athlete* or the Dutch *melak*. Syllable preference laws and universal tendencies make it possible to predict in what sort of environments epenthesis is more or less likely to occur (cf. Murray 1989). The problem is interesting, both from the viewpoint of the linguist looking at processes of language change and one who is interested in the history of particular changes in specific languages.

For an Indo-Europeanist, the difficulty of working with such problems is compounded by the fact that data for the older languages are written and one can never have recourse to calling up a helpful informant when the data are unclear. However, Indo-Europeanists must constantly reexamine the orthographic data and the clarification of the relationship between language perception by writers and inventors of scripts and the phonological systems of the languages in light of present day knowledge of how languages work.

In the following paper one such problem is examined, epenthesis in Classical Armenian, in terms of the script and the perception thereof. I hope to show that the system of the script has misled those dealing with the language itself.

Standard handbooks of Classical Armenian describe in greater or lesser detail vowel epenthesis, the insertion of a vowel in unetymological contexts. However, none of them encapture the rules and data systematically.² All treatments discuss epenthesis more or less as a dynamic process. Godel's (1975) discussion of Armenian syllable structure explicitly describes the epenthetic schwa as being underlyingly zero.³ Thomson (1989) contains the most extensive description of

¹I would like to thank Juliette Levin, Bill Darden, and an anonymous referee for their helpful comments and suggestions.

²In the following, this definition of epenthesis will apply, although the phenomena subsumed by the term include, e.g., prothesis, etc.

³One of T. Gamkrelidze's students is working on a thesis on the subject, but I have not yet been able establish contact with her.

the data, although he does not attempt a codification of rule systems. My own investigations have shown how to describe the attested patterns, but raise the more interesting question whether vowel epenthesis is a rule in Armenian or simply the lexicalization of language cooccurrence restrictions.

First, a brief note about the source of phonological data. Classical Armenian has been written since the 5th century A.D. in an alphabet ascribed to Maštoc' or Mesrop. Figure 1 shows the complete alphabet with a standardized transcription.

| | | | | | |
|---|---------|---|--------|---|-----------------------|
| Ա | a | Խ | x | Շ | č' /č'/ |
| Բ | b | Ծ | c /ts/ | Պ | p |
| Գ | g | Կ | k | Ջ | j /dž/ |
| Դ | d | Բ | h | Ղ | ř |
| Ե | e | Ճ | j /dz/ | Ս | s |
| Զ | z | Լ | l /l/ | Վ | v |
| Է | ē | Ճ | č | Ս | t |
| Ը | ə | Մ | m | Ր | r |
| Թ | t' /t'/ | Ե | y | Ց | c' /ts ^h / |
| Ժ | ž | Ա | n | Լ | w |
| Ի | i | Շ | š | Փ | p' /p'/ |
| Լ | l | Ո | o | Բ | k' /k'/ |

Figure 1. Armenian Alphabet and Transcription.

This alphabet is derived from Greek, although with many changes to fit it to the expanded phonological inventory of Armenian. It has 36 characters which represent both phonemes and numbers. The writing system has been adapted extremely well to Armenian and does not need to fall back on diacritics or digraphs to express the full range of Armenian vowels and consonants. All manuscripts are of later provenience, but the classical language lived on in literary and religious contexts long after evolving to Middle and Modern Armenian.

Knowledge of Classical Armenian vowel epenthesis is based on orthographic and grammatical peculiarities of the language, as well as the traditional pronunciation. In Greek inscriptions with Armenian proper names there is some evidence for the epenthesis (cf. Schmitt

1981). Although schwa-epenthesis is normally not expressed in Mesrop's orthography, when a word is divided in a manuscript so that a vowelless cluster of consonants might remain on one line, the scribes would use the sign <Է> to indicate the schwa's presence and quality. For example: *t' rč' unk'* but *t' ər-č' unk'* 'birds'. Grammatically, words which the unwritten epenthetic vowel renders disyllabic are handled differently from monosyllables. For example: *mnasc' es* [i.e., *mənasč' es*] 2nd sg. aor. subj. 'you might remain' where *-č' c'-* has dissimilated to *-sc'-*. (cf. Džaukjan 1982). In monosyllabic stems, there occurs no change, e.g., *c' uc' c' es* 2nd sg. aor. subj. 'you may show' (Meillet 1913:21). Similarly, the 3rd sg. aor. act. ind. of monosyllables shows the augment, where multisyllabic words do not. For example, *eber* 'he bore' but *gnac'* [i.e., *gənac'*] 'he went'. A study of actual manuscript practices in the oldest preserved documents would be extremely useful for separating Classical Armenian epenthesis from perhaps later overlays or interpretations based on a traditional pronunciation.

1.0 Rules of Epenthesis. Licit non-final syllable forms are CV(R)C (cf. Vennemann 1986:31 for Armenian syllabification). Possible syllable final clusters include: <rt, rt', nc', lč, nd, rg> from Thomson's examples, i.e., resonant + obstruent. However, as will be seen below, the list of possible syllable final clusters can be extended if only Wortauslaut is considered. Meillet (1936:52ff) points out that Classical Armenian tends to avoid consonant clusters, especially in syllable initial position. Of examples I have found, syllable initial clusters are practically non-existent. After syllabification (where vowels are syllabified, then leftward consonants, then rightward consonants) epenthesis occurs in the following manner to eliminate unsyllabifiable environments:

1. Ø > ə / C_ R'⁴

| | | | | |
|------------------|---|--------------|---------------|-----------------|
| ⁴ otn | > | otən | | 'foot' |
| dustr | > | dustər | | 'daughter' |
| inn | > | inən | | 'nine' |
| anxndroł | > | anxəndroł | | 'not demanding' |
| xorhrdoy | > | xorhərdoy | (cf. xorhurd) | 'mystery' gen. |
| ansrbut'iwn | > | ansərbut'iwn | (cf. surb) | 'impurity' |

2a. $\emptyset > \text{ə} / \#_ \{s/z\} 'C^5$

2b. $\emptyset > \text{ə} / \#_ \text{š}'t^6$

3. $\emptyset > \text{ə} / C' _ ^7$

($\langle x \rangle$ in rules indicates an as yet unsyllabified segment)

Use of feature notation will not affect this analysis. The environments for these rules are of increasing degree of consonantal strength (cf. Vennemann 1988:9; Berg 1989).

Note that the syllable may extend into the domain of other words in a sentence, thus applying post-lexically and effectively precluding the epenthetic vowel's arising. Thus *grel* > *gərel* 'to write' but *kare grel* 'he can write' with no schwa; or *mukn* > *mukən* 'mouse' but *mukn ase* 'the mouse speaks' (Jensen 1959:15). Given the nature of the orthographic system, one must query how to determine the classical pronunciation of these examples. Jensen does not give specific attestation proving his claim.

This analysis of the rules stems from work with Classical Armenian. In her 1985 MIT dissertation, Juliette Levin arrived essentially at the same set of rules for Modern West Armenian (p.336). My reaction was mixed. On the one hand I felt that I had reinvented the wheel, on the other hand, I was gratified that we reached essentially the same results independently. However, a third reaction was one of suspicion about such an epenthesis rule being preserved for so many centuries.

The following examples show words that would have an incorrect vowel insertion by the rules given: *anxlčmtank'* > **anxəlčəmtank'* 'lack of scruple' with incorrect schwa-insertion by rule 1 instead of the correct *anxəlčmətank'*. *manawand* 'especially' takes no epenthetic

| | | | |
|-----------------------|---|-------------|-------------|
| ⁵ zgal | > | əzgal | 'to feel' |
| sk'anč'eli | > | əsk'anč'eli | 'wonderful' |
| ⁶ štemaran | > | əštemaran | 'barn' |
| štapel | > | əštapel | 'to hasten' |
| ⁷ krak | > | kərak | 'fire' |
| srel | > | sərel | 'to cut' |
| šp'ot' | > | šəp'ot' | confusion' |

vowel, but *band* > *banəd* 'word', or *cnund* > *cənuṇd* and not **cənuṇəd*. Also, *mak's* 'excise tax' but *bans* > *banəs* 'words' and *sk'em* > *sək'em* 'monk's habit' but *sk'anč'eli* > *ək'anč'eli* 'magnificent'.

Two divergent problems bear on the discrepancies in vowel insertion above. First, some words or morphemes seeming to be exceptions to the rules are originally loanwords from Iranian or other languages. Armenian has so many loans from other languages that in the 19th century it was classified for some time as a Persian dialect (cf. Hübschmann 1897:xvi-xvii). Thus, for example, *štemaran* 'barn' comes from Persian. Investigations of the available works on Armenian etymology, primarily Meillet (1936) and Hübschmann (1897) failed to show any clear instances of native Armenian lexemes with initial *št*. The closeness of place of articulation of *š* and *t* do make them different than other clusters with *š* + occlusive.

By excluding words of foreign origin for the moment, most of the apparent exceptions can be eliminated. However, there is still the problem of *anxlčmtank'*. As noted in Thomson (1975:111), an epenthetic vowel often will appear in places where an underlying *i* or *u* is located. Thomson (1975:7-9) discusses at length the variations of vowel deletion. Leaving alone the various exceptions, the rule is:

4a. *i, u* > Ø / [- stress].

The numerous instances of /*i,u*/ in non-stressed environments which are evident in the table of examples are from diphthongs which have been reduced from underlying diphthongs.

I propose that instead of the traditional vowel deletion rule we substitute a vowel reduction rule:

4b. *i,u* > ə / [- stress]

One restriction would have to be made on the rule to delete the schwa if and only if doing so would not create unsyllabified elements.

5. ə > Ø / V(R)C_CV

The idea that *i* and *u* are reduced instead of being deleted is not new (cf., e.g., Winter 1962), but particularly in synchronic descriptions,

one finds reference again and again to a deletion. The analysis of words such as *anxlčmtank'* prove the reduction.

The underlying structure of *anxlčmtank'* must be *an-xilč-mit-an-k'* where there are attested the morphemes *xilč* and *mit*. By the application of Classical Armenian stress (construct R-dom word-tree, i.e., always stress the last syllable) we set up the environment for the reduction of the vowels and have the surface form *anxəlčmətank'* without a double deletion and epenthesis. *č'araxndrac'* 'of those who seek evil' (gen. sg.) could have an underlying structure *č'ara-x?ndir-ac'* > *č'arax?ndərac'* > *č'araxəndrac'*. (The question mark indicates my uncertainty as to the origin and underlying structure of the first syllabic element in *xndir*). Obviously, we can supply this vowel by the epenthesis rules, but that seems unnecessarily complicated.

This analysis can also take care of the discrepancies in dealing with the two plural markers *-k'* and *-s*. If one posits *-s* to be underlying *-əs*, then the so-called epenthesis is explained as a retention of the schwa. (The stress rule must now read: build Quantity Sensitive, R-dom word-tree because schwa is never stressed. Some schwas are thus underlyingly so whereas others are reduction vowels.) That Jensen (1959:15) and Thomson (1975:112) disagree on the pronunciation of words like *nawk'* 'ships' as either *nawək'* (Jensen) or *nawk'* (Thomson) may indicate the later change of the Classical Armenian diphthong *aw* to *av* and its effect on syllabification in traditional pronunciation. We definitely need a reexamination of the evidence in the manuscripts in order to solidify our views.

If the vowel reduction rule applies, then for a great number of the so-called instances of epenthesis no vowel was ever deleted and thus none must be inserted.

2.0 Epenthesis as Lexicalized. As noted above, many so-called epenthetic vowels are actually reduction vowels in non-stressed environments. Why, then, are they generally included in traditional grammars with the epenthetic vowels? I propose, that the writing system of Classical Armenian is to blame. Although Armenian possesses the sign <Է> representing /ə/, orthographic rules proscribe its use with two exceptions. First, in Anlaut of monosyllabic words and compounds with them; second, in word division at the end of a line to prevent a purely consonantal cluster (orthographically).

In the instance of reduction vowels one can thus characterize the

“epenthesis” as orthographic. The vowel has always been in place except in orthography. However, the reader or learner of written Classical (or Modern) Armenian must insert the reduced vowel according to his knowledge of actual pronunciation, or failing that, according to his knowledge of the lexicon. Murray (1989:295-296) notes that orthography can influence both a speaker’s output and his perception of his output.

I have already posited underlying schwa in the plural marker *-əs* and this raises the question whether the all other “epenthetic” vowels might not also be underlyingly present as schwa. Under this analysis, at one period a rule became valid for Classical Armenian inserting an epenthetic vowel in certain environments. Once this had occurred, the native learner acquired these words with the particular environmental cooccurrence restrictions as lexical items. Just as many modern Dutch speakers will learn the word for ‘milk’ as disyllabic *melək*, so too would the Armenian have more syllables in his speech than is indicated in the orthography. Under such an analysis, the forms listed in the footnotes as output of epenthesis rules are actually already the underlying forms. For example, the underlying form of *otn* ‘foot’ is *otən*.

A more general problem arises in that this analysis of epenthesis has implications for epenthesis in language description in general. Is epenthesis ever an active part of a synchronic grammar or is it always the lexicalization of syllable preference laws? Epenthesis may be present as an active rule if there is synchronic alternation within the language. For example, in some affective contexts, an American English word such as *incredible* shows an epenthetic vowel: *incəredible*. Another way to test whether epenthesis is active is to examine foreign borrowings in systems with regular but non-uniform epenthesis rules, i.e., where epenthesis occurs differently in a large array of environments. If borrowed words are adapted with general regularity to the rule system, then the epenthesis may still be in effect. If, on the other hand, the epenthesis rules are lexicalized fossils, then foreign words will be more or less well integrated into the system and may either violate some of the earlier cooccurrence restrictions or be adapted differently than words with the same environments of an earlier period. It may prove impossible to draw the line between genuine dynamic epenthesis and modification of new vocabulary on an ad hoc basis as needed.

The most general question this all raises is whether a generative rule is always necessary, even if it explains all data.

Why do some foreign words violate the rules? Foreign words which violated the "preferred syllable structure" of Armenian were adjusted according to available prototypes. In instances where the violation of syllable structure was against models of available words, but within the parameters of more universal syllable constraints, the foreign word could be adopted with an abnormal structure (e.g., *mak's*, *yašt*).⁸

Corroboration for this analysis is seen in the differing treatment of loanwords from native words. Thus, for example, *sk'em* > *sək'em* 'monk's habit' from Greek, but *sk'anč'eli* > *ask'anč'eli* 'magnificent' (native) shows that rule 2a is older and affects native vocabulary. After the environment for the rule was eliminated, the rule quit applying, i.e., it bled itself to death. When later words were borrowed which reinstalled the environment, the more general rule 3 applied. Whether it applied as an active dynamic rule or rather as a prototypical generalization is unclear, I tend towards the latter. *štemaran* was borrowed from Iranian which has an initial epenthetic vowel so that the borrowing may have taken this over although not writing it. In the case of, e.g., *yašt*, the syllable final restriction was only valid for a limited period and later words could be assimilated into Armenian vocabulary without deformation, even though there were until then no native models.

The adaptation of foreign words with otherwise illicit clusters in Auslaut could explain the above mentioned fact that in Armenian there is greater variety in word final consonant clusters than syllable finally in other parts of words. However, there are many native words which similarly end in consonant clusters that are illegal elsewhere. For example: *əst* 'in,' *ost* 'branch.' An explanation for this environmental restriction may be found in the relative chronology of sound changes. As has been mentioned above, Classical Armenian has word final accent except if the final vowel is schwa. Knowledge of the *Auslautgesetze* shows that this accent was previously on the

⁸Cf. here the cooccurrence restriction of non-homorganic voiced stops in IE which may be a relic of an earlier cooccurrence restriction on glottalized stops. The later pattern is inexplicable except as a retention of earlier patterns despite loss of justification. This is an area in which to tread warily when setting up rule systems, for the current rule may be a lexicalization of a previous rule's output.

penultimate syllable (i.e., construct binary, left dominant feet and a right dominant word-tree, or final syllable is extrametrical and build right dominant word-tree). If the cluster simplification process is prior to the loss of final syllables, this will account for the illegal clusters. At the time of the epenthesis, they were followed by vowels and the consonants were not in the same syllable. For example, the word *óst* 'branch' derives from **óstos* in which the cluster is perfectly acceptable.

Another problem which may be clarified by this view of cluster development is the perennial question of Armenian *k'san* 'twenty'. The traditional explanation has an IE **wīkṛti* or the like develop into Proto-Armenian **gisán*. Deletion of /i/ leads to the form **gsán* after which occur devoicing and aspiration. Winter (1965:106-107) has pointed out that the Classical pronunciation *k'əsan* would make the devoicing rule unlikely. He and others have tried to explain the form in terms of laryngeals. Recently, Huld (1984) has proposed a reanalysis of the form in sandhi as the cause of the odd phonological form. If application of cluster cooccurrence restrictions came after the apocope, then the devoicing is explained. Then clusters are broken up by the epenthesis rule. Note here that a conflict arises with the above analysis of vowel reduction. However, assuming that the vowel reduction is in those locations where a lexical form is available with the full vowel, the conflict is resolved. In fact, justification for the original epenthesis rule may be sought in the deletion of vowels which created unacceptable consonant clusters in words with available non-deleted full forms. Here the proposed developments:

**wīkṛti* > **gisánX* > **gsánX* > **k' sánX* > *k'əsan*⁹

To justify this analysis of Classical Armenian, a precise study of manuscript material is in order as well as the patterns of borrowing from other languages. Hübschmann (1897:14) has very interesting material in this respect on the chronology of loanwords from Iranian and their differing treatment in Armenian. Iranian loans come in predominantly during the Parthian period but also later during Sassanidian rule. Words borrowed during the earlier period conform

⁹The X indicates some indefinite element because the question of the relative chronology of the Auslautgesetze is a matter for a different paper.

to Armenian rules of vowel reduction (or deletion) in the unstressed syllables just like the native words. For example, the Iranian *gund-* is borrowed in an early compound Arm. *gndapet* 'army leader' where *u* > ə. In a later borrowing with the same first element, the *u* is maintained as such: *gundsatar*. The above-mentioned Greek loanword of the 4-5th centuries AD. *sək'em* < ΣXHMA 'monk's habit' has a different vowel insertion than the native *əsk'ančeli* 'wonderful' showing that rule 2a may have died out before the Greek word was borrowed.

An /i/ in early loanwords behaves just like native /i/. Later loans often maintain the /i/ of the original language even in the unstressed position. It would be interesting to know if there are instances where the unstressed, maintained /i/ of a non-final syllable in foreign word becomes stressed and changes to /ē/ in analogical inversion of the rule for native vocabulary:

$$\bar{e} (< *ei) > i \text{ / [- stress]}$$

This rule and a similar one explain the occurrences of unstressed *i* and *u* alluded to after rule 4.a. Examples of these rules include: nom. sg. *mēg* 'fog' vs. gen. sg. *migi* and nom. sg. *sēr* 'love' vs. gen. sg. *siroy*; *loys* 'light' vs. *lusawor* 'bright'.

Exceptions of later borrowed foreign words to the rules of epenthesis indicate that the rules are no longer valid as dynamic processes but rather that borrowings are adapted in an ad hoc manner to the pre-existing canonical restrictions on syllable structure.

3.0 Conclusions. The conclusion that Classical Armenian Vowel Epenthesis is lexical has ramifications for numerous other analyses in various languages. For example, one can argue that the English rule of consonant assimilation in the Latin privative prefix *in-* is also purely lexical. This is a textbook example of consonant assimilation. The typical rule will state that *in-* will assimilate to the following consonant so that we have: *impossible*, *irrevocable*, *illegal*, *inoperable*. However, the forms with the native prefix *in-* do not undergo assimilation: *inlet*, *inlying* etc. Solutions to the problem include labelling this particular prefix as undergoing the rule, using levels in lexical phonology, or positing an underlying structure of /iN-/, i.e., stating that the resonant is undefined, taking on the features of what follows or defaulting to *in-*. However, the set of forms with variation dates back to Latin, and

many made their way into English through French. Just as the retention of the epenthesis rule from Classical to Modern Armenian (cf. above on Levin 1985) makes me uncomfortable, so too does the active retention of this assimilation rule. A number of informants have intimated that they feel the Latinate variants to be separate prefixes: cf. *obvious* : *?obvious* where one may elicit *ilobvious*, although this runs counter to the underlying structure generally assumed. My guess is that there is a set of negative prefixes (*il*, *in*, *ir*, *im*) which are used in phonologically determined environments. An early change is preserved in the surface forms of English just as in Armenian with vowel epenthesis.

Lack of synchronic variation of epenthesis, varying application of reduction and epenthesis rules for native and loanwords as well as for different chronological layers of loanwords, and the relative chronology (of reduction, assimilation, loss of final syllables, and epenthesis) all suggest the strong probability that Classical Armenian vowel epenthesis is a metalinguistic device to describe lexicalized syllable preference laws. The native writing system may have helped cause the confusion because it requires the non-native speaker (i.e., everyone) to learn vowel insertion rules. These rules are not a reflection of dynamic rules.

BIBLIOGRAPHY

- Berg, Thomas. 1989. "Intersegmental Cohesiveness." *Folia Linguistica* XXIII/3-4: 245-280.
- Džaukjan, G. B. 1982. *Sravnitel'naja grammatika armjanskogo jazyka*. Erevan: AN ASSR.
- Godel, Robert. 1975. *An Introduction to the Study of Classical Armenian*. Wiesbaden: Reichert.
- Hübschmann, H. 1897. *Armenische Grammatik*. Leipzig: Breitkopf & Härtel.
- Huld, Martin. 1984. "On Armenian k'san." *Annual of Armenian Linguistics* 5:59-67.
- Jensen, Hans. 1959. *Altarmenische Grammatik*. Heidelberg: Winter.
- Levin, Juliette. 1985. A Metrical Theory of Syllabicity. Ph.D. diss. MIT. Cambridge.
- Meillet, Antoine. 1913. *Altarmenisches Elementarbuch*. Heidelberg: Winter.
- _____. 1936. *Esquisse d'une grammaire comparée de l'arménien classique*. Wien: Mechitaristen.
- Murray, Robert W. 1989. "On Epenthesis." *Folia Linguistica* XXIII/3-4: 293-316.
- Schmitt, Rüdiger. 1981. *Grammatik des Klassisch-Armenischen*. Innsbruck: Institut für Sprachwissenschaft.
- Thomson, Robert W. 1975. *An Introduction to Classical Armenian*. Delmar:

Caravan.

_____. 1989. *An Introduction to Classical Armenian*. 2nd Revised Edition. Delmar: Caravan.

Vennemann, Theo. 1986. "Syllable-Based Sound Changes in Early-Armenian." *Annual of Armenian Linguistics* 7: 27-43.

_____. 1988. *Preference Laws in Syllable Structure*. Berlin: Mouton de Gruyter.

Winter, Werner. 1962. "Problems of Armenian phonology III." *Language* 38/3(1): 254-262.

_____. 1965. "Evidence in Armenian." in Werner Winter (ed.). *Evidence for Laryngeals*. The Hague: Mouton, 100-115.

A FEMININE/DIMINUTIVE SUFFIX IN EARLY OSSETIAN¹

David Testen
University of Chicago

Lacking gender as a grammatical category, the modern Ossetian language has come to lose the numerous morphological mechanisms which provide the Indo-European languages with the means to derive feminine nouns.² In part, this loss of gender may be ascribed to the geographical location of Ossetian in the heart of the Caucasus, where it is surrounded by languages such as Georgian, in which gender plays no grammatical role. The areal factor notwithstanding, it should not be forgotten that internal developments within Ossetian itself no doubt tended to place gender in jeopardy: the phonological erosion of the word-final syllable must have imperiled the prime locus for the inherited Indo-European gender-marking elements, while the drastic reduction of the case system, and its subsequent replacement by a new, agglutinative paradigm, must have contributed critically to the erasing of much of the morphological substance through which the category was expressed.

It will be the purpose of the following pages, however, to suggest that the loss of gender—or, to be more specific, the loss of the derivational mechanisms for marking the feminine gender—was not simply a matter of an inexorable plunge into oblivion. An investigation of the lexicon of the modern language, we shall seek to demonstrate, provides evidence that a secondary system of suffixation marking many feminine entities developed on the ruins of the Old Iranian system, and that this secondary system had much in common with the derivational suffixes of certain Middle Iranian languages, such as Sogdian. The reconstructed pre-Ossetian suffix is also associated with a set of forms which suggest that it came to acquire an expressive (diminutive or pejorative) sense. Finally, we shall look into the possibility that the system of suffixation reconstructed here may cast some light on the scant linguistic evidence available concerning the

¹The author thanks Roland Bielmeier, who served as referee for an earlier draft of this paper, for his valuable comments. All errors remain the responsibility of the author.

²Animate nouns may be overtly marked as referring to male or female entities by incorporating them into a compound whose first element is *næl*- 'male' (< Iran. **narya*-) or *syl*- (< **sy*- < Iran. **stri*-)—cf. *næl xæræg* 'male donkey,' *næl biræγ* 'male wolf,' *syl xæræg* 'female donkey,' *syl biræγ* 'female wolf.'

ancestors of the Ossetians, the Scythians and Sarmatians of the classical sources and the Alans of the medieval sources.

Of particular importance for this problem is Ossetian *čynz* (in the western dialect, Digor, *kinzæ*) 'bride, daughter-in-law.' The Indo-Iranian pedigree of the stem of this form is beyond doubt—cf. Sanskrit *kanā-* and *kanyā-* and Avestan *kainiiā-* 'unverheiratetes Mädchen.' Abaev (1958: 607) reconstructs the ancestor of the Ossetian form as **kantī-* (*kan* + *tī*), stating that the ending *-tī* is characteristic of the East Iranian languages (cf. Sogdian *knc* 'girl' and, from the Pamir group, Wakhan *könd*, *kənd* 'woman, wife' and Munjan *kintika*).

The suffix **-tī-* which Abaev posits for this word must be regarded as highly conjectural, however. While there can be no doubt that an Ossetian affricate may result from early Iranian **-ti-* (cf. *balc* 'journey' < **bār-ti-*, (Iron) *ssæz* 'twenty' < **vinsati-*, etc.), this sequence is only one of several possible sources for Ossetian *c* and *z*. Most typically, such affricates have developed from Iranian **č* and **j*, the palatalized counterparts of the velars **k* and **g* which were phonemicized at an early stage of Indo-Iranian. In addition to **č* and **j*, Ossetian affricates have been traced back to **θy* (as in *æcæg* 'true' < **haθya-ka-*) and, according to Gershevitch, to **xš* (if his discussion (1985: 127-138) of *wac* 'spirit' and similar forms is correct). Since a proto-form **kantī-* affords no clear morphological explanation for the ending **-tī-*, we should attempt to determine whether an alternate analysis of the *z* of *čynz* is possible whereby the morphological constitution of this word is more readily comprehensible.

In connection with the *z* of *čynz*, we may adduce Benveniste's comments on the anomalous suffix *-zæg*.

La remarque [concerning the isolated nature of the suffix *-iwæg* (Digor *-ewæg*)—DT] vaut aussi pour un suffixe dont la réalité même a semblé douteuse à V. Miller, le suffixe *-dzæg*, qu'il ne mentionne qu'avec un point d'interrogation. Il n'en donne que deux exemples, qui en effet ne sont probants ni l'un ni l'autre: dans *ualdzæg* "printemps", seul l'étymologiste peut isoler *-dzæg* en comparant *ual-* à av. *vaŋri* "au printemps"; et *qældzæg*, d. *i γældzæg* "alerte, joyeux; plaisir"... n'est pas en relation de sens évidente avec *qal*, d. *i γal* "éveillé; éveil". Cependant nous en avons rencontré un exemple meilleur: *uæmdzæg* "vomissement" (*uæm-in* "vomir"). Il faudrait colliger d'autres exemples, s'il en est, et dégager la fonction du suffixe avant de commencer seulement à s'interroger

sur la préhistoire de *-dzæg*, qui matériellement suppose **-čaka-*. Ici le travail préliminaire reste à faire. (1959: 110).

Benveniste is quite correct in reconstructing this ending as **-čaka-*, but perhaps the most interesting feature of the suffix *-zæg* is its predilection for stems ending in *-l*. The examples which he and Miller adduce cannot be separated from certain other cases of forms containing *-lʒ-* concerning which Benveniste makes the following observations:

Il y a enfin un groupe oss. *-ldz-* qui apparaît à la jonction du thème et du suffixe dans quelques mots: *ænguldzæ* "doigt" [the Digor form; the Iron counterpart is *æng*ylʒ*—DT]; *mældzīg*, d. *muldzug* "fourmi"; *ualdzæg* "printemps". Il faut, à notre avis, interpréter ici le groupe *-ldz-* en donnant à *l* la valeur "étymologique" de *-ri-* final de thème, précédant un suffix *-č-*. On peut montrer que dans les trois mots cette restitution phonétique est satisfaisante pour la morphologie (1959: 32).

Benveniste astutely points out that the phonetic shape of *ængulʒæ*, *mælʒyg/mulʒug* and *walʒæg* call for a vowel **i* to have originally occurred at the end of the stem, since the presence of *l* rather than **r* would otherwise be inexplicable. As he indicates, this assumption is borne out by comparative evidence: *ængulʒæ* < **anguri-ča-* (cf. Sanskrit *anguli-*, *anguri-*), *mælʒyg/mulʒug* < **marvi-ča-* (the details of the development remain obscure in many respects, but cf. Avestan *maoiruui-*, OCS *mravijь*), *walʒæg* < **vāri-ča-* < **vahri-* (Benveniste 1959: 32).

However, Benveniste does not note the equally conspicuous corollary to these data, viz. the clear connection between the *i*-final stems which he reconstructs and the occurrence of the infrequent suffixes beginning with *-ʒ-*. It is not difficult to imagine that there could be a causal relation between these two facts, if we posit that the Iranian nominal suffix **-ka-/*-kā-*, reflexes of which (*-æg*, *-ag*, etc.) are so familiar throughout the Ossetian lexicon, was palatalized to an affricate when preceded by a stem ending in *-i-*. Such a palatalization would have had much the same effect as the so-called "third palatalization" of Slavic, which likewise created affricated realizations of the Slavic reflex of the **-k*-suffix:

| | | | |
|--------------|------------------------|----------------------|----------------------|
| | <i>æng*ylʒ</i> | < <i>*anguri-ča-</i> | < <i>*anguri-ka-</i> |
| like Slavic: | <i>*otьcb</i> 'father' | < <i>*oti-cu</i> | < <i>*oti-ko-</i> |

This hypothesis makes available another interpretation for *čynz*. Since comparative data make it clear that there was originally an *i*-class noun lying at the heart of this form, it is not difficult to imagine that it is the **k*-suffix, rather than Abaev's hypothetical **-tī-*, that played a role in the development of Ossetian *čynz*; such a suffix is directly attested elsewhere in Iranian for this noun (cf. Avestan *kainikā* 'Mädchen,' Middle Persian *knyk*). *Čynz* could therefore have developed regularly from **kani-kā-* if we assume (a) the palatalization of **k* posited above, (b) the umlaut of the stem-vowel under the influence of the vowel **i* (which is a well-documented phenomenon for Ossetian, although the specific factors which govern its appearance have yet to be pinned down), and (c) the subsequent elision of the vowel **i*, as we have had to posit for *æng^wylz*, *walzæg*, etc.³

The assumption that the suffix **-ka/*-kā*, when added to an **i*-stem noun, led to the appearance of an affricate enables us to fill in a lacuna observed among the suffixes of Ossetian. The modern endings *-æg* and *-yg* (= Digor *-ug*) clearly show the results of adding the **-k*-suffix to an **a*-stem and a **u*-stem, respectively (cf. (Digor) *næwæg* 'new' < **nava-ka-*, *bazyg* (Digor *bazug*) 'shoulderbone' < **bāzu-* 'arm' + **ka-*); at least some instances of the ending *-ag* have arisen from consonant-stems, presumably from the Old Iranian nominative singular form—cf. Ossetian *fændag* 'road' < Iranian **pantah-*, **paθah-* (nom. sing. **pantāh*), Ossetian *dændag* 'tooth' < **dantan-* (nom. sing. **dantā*). There is no sign, however, of an ending **-yg* (Iron)/**-ig* (Digor) which we might have anticipated as the result of adding **-ka/*-kā* to a stem in **-i*. This absence may be explained if we assume that this sequence underwent a palatalization of the sort mentioned above.⁴

A close parallel to the palatalizing effect of a preceding **-i-* which

³This interpretation of *čynz* is also found expressed in Bielmeyer (1993: 7).

⁴There do, however, exist instances of an ending **-ikk-* (Iron *-ykk*, Digor *-ikkæ*), with gemination, which apparently conveys a diminutive sense—cf. *wærykk* 'lamb,' *særykk* 'kid,' *becykk* 'forelock,' *mætyykk* 'type of plant.' The historical source of Ossetian geminated stops remains unclear in many respects. Nevertheless, one cannot rule out the possibility of an earlier affricate stage for **-ikk-*, since it is likely that the ordinals *dykkag/dukkag* 'second' and *ætykkag/ætiikkag* 'third' passed through such a development (< **duč-*, **θrič-* < Old Iranian **dvitiya-*, **θritiya-*)—cf. Ossetian *dyccæg* 'Tuesday,' *ætyccæg* 'Wednesday.' It is quite conceivable, therefore, that the ending **-ikk-* may constitute another manifestation of the **-ič-* suffix postulated here.

we posit for early Ossetian may be seen in Sogdian. Gershevitch cites several instances of Sogdian \check{c} < **k* "... [p]alatalized by a preceding *i* which afterwards usually disappears..." Among them are **knc* 'girl' < **kanikā*-, *zm'wrc* 'ant' < **marvika*-, and *nyc* 'nostril' < **nāhikā*- (1954: 39).

This discussion of *čynz* might conclude here as an etymological detail of rather limited significance were it not for the fact that the Old Iranian *i*-class consisted of feminine nouns, largely derived from consonant-stems (cf. Avestan *ašaon-i*-, f. of *ašaunu-* 'pious'), from *u*-stems (Avestan *pərəθuuī*-, f. of *pərəθu-* 'broad') and occasionally from thematic stems (*daēuu-i*-, f. of *daēuu-a-* 'demon') (Reichelt 1978: 162). We may easily imagine the significance which a palatalization of **-k-* > **-č-* would have acquired as the loss of the distinctions in word-final syllables threatened the conventional markers of gender in the language: once the Old Iranian feminine markers such **-ā*, **-ī*, etc., disappeared, the new suffix **č* would have survived as virtually the sole marker which was unambiguously feminine. Adding the **-k-* suffix to a substantive thus provided a means for maintaining earlier oppositions, reflected now not in the nature of the vowel but in the nature (palatalized vs. non-palatalized) of the consonant. Under the circumstances, it would not be surprising to find that the palatalized suffix came to be used even in places in which it was not etymologically justified (i.e., where it was not originally preceded by a palatalizing **i*).

A development of this sort evidently took place in the history of Sogdian, where it led to the appearance of an opposition in adjectival suffixes composed of masculine *-y(y)* (< **-aka-*) vs. feminine *-c(')* (< **-ičā-*)—cf. 'wswyc, 'wswxtc 'pure (f.)' vs. 'wswɣtyy 'pure (m.),' 'ktc' 'done (f.)' vs. 'ktyy 'done (m.)'. Note that the opposition between the masculine reflex of **k* and the feminine reflex of **č* has been expanded beyond the original scope of *-č* in etymological terms: both 'wswxtc and 'ktc', for example, are thematic participles which in Old Iranian must have had feminine derivatives in **-ā-* (**ava-suxtā-*, **kṛ-tā-*), and for which we would therefore not anticipate the ending *-c(')*.

Comparable data may be adduced from Khwarezmian: *wdnc* 'old (f.)' vs. masc. *wdnyk*, *γrmncy* 'hot (f.)' vs. masc. *γrmnd*, 'mc 'dead (f.)' vs. masc. *mdyk* (Humbach 1989: 198).⁵

⁵Morgenstierne (1973: 102-107) challenges Gershevitch's derivation of Sogdian

It appears likely that a similar expansion took place in early Ossetian. As we have seen, gender is no longer relevant to the grammatical system of the modern language. There are, however, a fair number of forms for which we may reconstruct an earlier ending *-č (frequently with *i*-umlaut of the stem vowel). Several, like čynz, have a clear feminine reference; others appear to be associated with small or diminutive entities, suggesting that the putative suffix *-č came to play an expressive role at some stage of the language. In several cases we find that the remnants of this suffix in modern Ossetian are no longer obvious, having melded with the stem to form a new, unanalyzable unit. Assuming the presence of this earlier *-č in fact provides a means of explaining a number of forms which might otherwise remain etymologically opaque.

One such form is *idæz* 'widow,' which is clearly to be related to Indo-Iranian **vidhavā-* and to its well-documented cognates elsewhere in Indo-European (Gothic *widuwō*, Old Church Slavonic *vdova*, Old Irish *fedb*, etc.), although the details of the development remain unclear. The final affricate of this word is said by Abaev (1958: 539) and Bielmeier (1977: 234) to be due to a suffix *-ti- like the one which was assumed for čynz, but since such a suffix is no more motivated in this case than in the case of čynz it is equally possible that we have here an instance of *-č appended to a stem **vidā-*, ultimately from Iran. **vid(a)vā-* (cf. the Sogdian adjective in *wydwč* 'ync 'widow woman').⁶

-c- via regressive palatalization of the velar, proposing instead an Iranian feminine ending *-ačī (< *-akī) corresponding to masculine *-aka-, and comparable to the *deva-/devī-* distinction of Indic. No such opposition is documented for the Old Iranian languages, however, and even in Indic the use of -akī- is limited. The question of the phonological development of the Ossetian etyma under discussion here is closely related to similar phonological questions elsewhere in East Iranian which have yet to be resolved. Morgenstierne points out that the distinction between, e.g., Shughni *vuđj* 'been (masc.),' and feminine *vic* cannot be ascribed to **būtaka-* vs. **būtākā-* (the latter of which would have given feminine **vađj*), but it is not clear that **būtičā-* (< **būtikā-*) is to be ruled out as a reconstruction in favor of **būtačī-*. The resolution of this and similar questions will be of considerable relevance to the Ossetian problems considered here.

⁶The survival of the initial syllable of **vid(a)vā-* in Iron *idæz* is curious, since elsewhere **vi-* disappears in this dialect (cf. Iron *ssæz* 'twenty' vs. Digor *insæj* < **vinsati-*, Iron *g*yr-yn* 'give birth' vs. Digor *igur-un* < **vi-kur-*). If our suspicions about the development of Ossetian *us* (Digor *wosæ*) 'woman' < **vawd-* < **vadū-*

Abaev has also identified a set of nouns which led him to assume the earlier existence of a nominal suffix **-tæ*. The forms which he cites—*cæst* (Digor *cæstæ*) ‘eye,’ *myst* (Digor *mistæ*) ‘mouse,’ *syst* (Digor *sistæ*) ‘spider,’ *nus* (Digor *nostæ*) ‘bride’—led him to conclude that the plural marker *-tæ* of the modern language originally served a more general purpose of conveying abstraction, generalization, collectivity, etc. (Abaev 1949: 572). If we bear in mind, however, that Ossetian *-st-* may develop regularly from **-sč-* (cf. *fæstæ* ‘after’ < **pas-ča*, *isty* ‘anything’ < **es-či*) it is possible to interpret these forms as consisting of an inherited stem ending in a sibilant and the suffix **-č*.

Iranian **čašman-* ‘eye’ (cf. Avestan *čašman-*) > Pre-Oss. **čašm* (cf. Iron *casm* ‘loop’; Digor *cans* ‘window opening’); **čašm-* + **-č-* > **čaš(m)-č-* > **čas-t-* > *cæst/cæstæ*

Iranian **muš-* ‘mouse’ (cf. Sanskrit *muṣ-*, NPers. *muš*) > Pre-Oss. **muš*; **muš-* + **-č-* > **miš-č-* > *myst/mistæ* (compare the change of the stem vowel in Digor to the umlaut seen in *čynž/kinžæ*)

Iranian **spiš-* ‘small insect’ (cf. Avestan *spiš-* ‘Laus, Milbe (od. dgl.)’) > Pre-Oss. **siš*; **siš-* + **-č-* > **siš-č-* > *syst/sistæ*

Benveniste (1959: 8) cites a phonological parallel to *cæst/cæstæ* which may well belong in this list, *fist/festæ* ‘wool from the spring shearing’ alongside *fasm/fans* ‘wool from the autumn shearing’ < Iranian **pašman-* (cf. NPers. *pašm* ‘wool’).⁷ As in the case of *myst/mistæ*, we find the effect of an *i*-umlaut in *fist/festæ*, which Benveniste ascribes to an ancient verbal noun suffix **-ti-* (i.e., **paš-ti-* ‘a shearing’ > **paišti-* > *fist/festæ*), but which could equally well be due to **-č-* < **ika-*.

Abaev (1949: 574) has also identified a set of forms which provide an indirect indication of an earlier alternation between **-k* and **-č*. In are correct (see below), is it possible that an analogous development gave rise to **viwd-* < **vidv-*, and ultimately, via assimilation of the semivowel (**(v)ijd-*), to *id-*?

⁷Abaev (1958: 476) observes that *fasm/fans* is “... of higher quality than *fist*...” The derivative *fist/festæ* (< **fasm-ča*) might therefore represent some form of pejorative expressiveness.

discussing the now unproductive suffix *-ynz/-inzæ*, for which he reconstructs an apparent diminutive significance, Abaev notes that this suffix is evidently close to the suffix *-yng/-ingæ* found frozen in a small number of nouns, and suggests that "... If *-yng* is to be traced to *-aina* + *-ka*, then *-ynz* can go back to *-aina* + *-ti* :

Suffix *-ynz/-inzæ*:

ærcynz/ærcinzæ 'needle for stitching shoes'; cf. *arc* 'spear' (< Iran. **ršti-*
wadynz 'reed-pipe'; cf. (prob.) *wad* 'wind' (< Iran. **vāta-*)
qædynz/γædinzæ (*γæzinzæ*) 'onion'; cf. *qæd/γædæ* 'stalk'⁸

Suffix *-yng/-ingæ*:

ruzyng 'window' < Iran. **raučah-* 'light'
talyng/talingæ 'dark (adj.), darkness'; cf. *tar* 'dark' (< Iran. **tanθra-*)
aryng 'dough trough'; cf. Iran. **ar-* 'grind,' **arta-* 'flour'

Once again, we may note that Abaev's posited suffix **-ti* has no obvious morphological motivation. The suffixal pair **-ingæ* and **-inzæ* are explained automatically, however, if we assume an earlier stage at which **-k* and **-č* alternated along the lines which we have conjectured: given a primary suffix **-in* < **-aina-*, the superimposing of the suffixes **-k* and **-č* would lead to **-in-k* and **-in-č-* respectively, the forms which Abaev reconstructed.

Two possible instances of a feminine suffix **-č* which are of particular interest may be found in the pair *æfsin* (Digor *æfsinæ*) and *æxsin* (Digor *æxsinaæ*). Both terms are used of a woman enjoying a respected position, and *æfsin(æ)* was long believed to be a synonymous variant of *æxsin(æ)*. Benveniste, however, observed that the two terms must be regarded as distinct.

L'origine de *æxsin* est dans **xšaiθnī-* (av. *xšōiθnī-*). En fait, le seul

⁸Benveniste, however, takes *qædynz* to have arisen from **gandačī-* 'strong smelling' under the influence of the nouns in *-ynz/inzæ* (1959: 16). It could, however, also have come about by dissimilation from **gandan-č-*, using the feminine suffix in question—cf. Persian *gandanā* 'leek.'

emploi *spécifique* que je connais de *æxsin* est dans la veille dénomination *uad-æxsin*, dig. -æ “maitresse des vents” (personnage de la mythologie populaire), où *æxsin* reste la qualification noble que l’ancien *xšaiθnī-* était déjà pour des déesses ou des femmes de haut rang.

Mais *æfsin* est tout autre chose. C’est un terme de relations humaines, l’appellation que la jeune mariée donne à la mère de son mari quand elle prend place dans sa nouvelle demeure, et aussi le terme de politesse pour les invités s’adressant à la maitresse de maison. Nous dériverons *æfsin* de **abi-šaiθnī-*, reconstruction dont l’avestique nous offre tous les éléments... Du thème *abi-šažtan-* déjà admis par Bartholomae, on tire un féminin **abi-šaiθnī* “résidente; maitresse de maison”, d’où régulièrement oss. *æfsin*. Le sens de l’appellation s’accorde bien à son emploi propre. La jeune mariée (*čindz*, d. *kindzæ*) s’adresse par ce terme à la maitresse du foyer où elle vivra désormais, tout comme elle appelle *xicau* “maitre” le père de son mari. Il est aisé de comprendre que *æfsin*, terme de respect, soit devenu rapidement, dans ces conditions, l’équivalent de *æxsin*, qui qualifiait les femmes nobles, et que ces appellations puissent être prises l’une pour l’autre. L’ossète a donc hérité de deux désignations indépendantes: **xšaiθnī-* “auguste”, oss. *æxsin*; et **abi-šaiθnī* “résidente; maitresse de maison”, oss. *æfsin*. (1959: 19-20)

Benveniste’s proposal is attractive, but the suffix **-č* which we have proposed affords an alternative explanation for *æfsin* lying closer at hand.⁹ It is clear that under certain conditions the **-č* in final position could become a semivowel *j*: cf. the numeral ‘twenty’ (Digor *insæj* vs. Iron *ssæʒ*), or the third-person singular present indicative ending *-j* < **-č* < **-ti*. When preceded by *n*, *j* is regularly metathesized to yield *-jn-* (the *n* of which is lost in the Iron dialect in word-final position after *o*—cf. Iron *æncoj* ‘peace’ vs. Digor *æncojnæ*).¹⁰

⁹If *æfsin(æ)* is from **abi-šaiθnī-*, as Benveniste proposes, we might expect a Digor form with initial *e-*, as is found in *ævjid/evged* ‘guarantee’ < **abi-kaita-* (Benveniste 1959: 13, 48).

¹⁰Roland Bielmeier (personal communication) has observed that **-nti* is retained as *-nc* in the third-person plural ending of the verb—cf. *færsync* (Digor *færsuncæ*) ‘they ask’—and that we therefore have no evidence that **-nčl-nc* developed into **-nj*. While it is true that *-ync/-uncæ* appears to entail such a development, the ending *-oj* (Digor *-oncæ*) of the third-person plural of the transitive past (*bafarstoj/bafarstoncæ* ‘they asked’) and of the subjunctive *færsoj/færsoncæ* ‘they will ask,’ both of which presumably derive from **-ānti*, suggests that the issue is more complex. I am more inclined to regard the development of the present indicative ending of *færsync/færsuncæ* (with its curious lack of voicing in a non-initial obstruent) as exceptional, and tentatively suggest that this absence of voicing

It is therefore entirely possible that a word ending with *-īn/-īnæ* may be analyzed as a development of **-in-č*, in which **-č* > **-j* and **-nj* > **-jn*. If we bear in mind that the feminine suffix **-č* which we have posited had the capacity for umlauting a preceding vowel, we must allow for the possibility that the **i* of the putative **-in-č* may be secondary. In light of these observations, it is worth noting Benveniste's characterization of *æfsin/æfsinæ* as "le terme de politesse pour les invités s'adressant à la maîtresse de maison." In the sense of 'hostess' *æfsin(æ)* is the feminine counterpart of *fysym* (Digor *fusum*) 'host' < Iranian **pasu-mant-* (lit. 'possessing cattle,' **pasu-*), and it is not difficult to imagine that the former term could be related to the latter by means of the addition of the feminine suffix which we have posited to **fsum*.¹¹ An assimilation of **-m* > **-n* under the influence of a following **-č* would be entirely natural, and the fronting of **u* > **i* would be paralleled by, e.g., our interpretation of *mistæ* 'mouse' < **muš-č-*:

**fsum-* 'host' (< Old Iranian **fšu-mant-*) (replaced by **fusum* < **pasum-*)

**fsum-č-* 'hostess' > **fšim-č-* > **fšinč-* > **fsinj-* > **fsijn-* > *æfsin/æfsinæ*

It is quite possible, in short, that Benveniste is correct in maintaining a distinction between *æfsin(æ)* and *æxsin(æ)*, but that *æfsin(æ)* is not to be taken as an isolated survival from Old Iranian but rather as a

has blocked the further development into *j*—a development which is seen in the case of the Iron third-person plural ending *-oj*. Since the preterite is formed periphrastically with an auxiliary 'be,' I would ascribe the Digor ending *-onca* to an innovative restructuring based upon the present indicative *ænca* 'they are' < Iranian **hanti*—contrast the Iron counterpart *ysty* 'they are'—the new ending *-onca* having expanded from there to the subjunctive. It must be granted, however, that much about the development of final **-č(a)* remains to be explained.

¹¹We note that *æfsin(æ)* is evidently to be traced back to **fsum-* < Iran. **fšu-mant-*, with the zero-grade of the first element, rather than to the **pasu-mant-* which the masculine *fysym/fusum* evidently entails. The shape **fšumant-* is supported by Avestan *fšumant-* and Sanskrit *kṣumant-*; the form *fysym/fusum*, presumably resulting from **fasum-* by umlaut, is due either to an alternate vowel-grade of the word (cf. Sanskrit *paśumant-*) or a restoration of the vowel of the first element of the compound on the basis of *fys/fus* 'sheep' < **pasu-*.

productive derivation dating from an intermediate stage of the language, the transparency of the derivation having been lost through an accretion of subsequent phonological developments. As for *æxsin(æ)*, Benveniste may well be right in tracing this form back to Old Iranian **xšaiθnī*- 'radiant'; alternatively, however, we cannot rule out the possibility that here too we find a remnant of the suffix *-č which has been rendered opaque over time. Under the circumstances, it may not be outlandish to note the existence of Christian Sogdian *xšwnč* 'queen,' the feminine counterpart of *xšwny* 'king' < **xšayvan-aka-*, which matches Ossetian *æxsin(æ)* semantically and constitutes a clear possibility formally.

A word such as *æfsin* demonstrates that the sound changes of Ossetian may give rise to forms in which the initial source is no longer at all obvious. In positing a suffix such as *-č, however, we cannot overlook the possibility that the (morpho-)phonological sequences in which this suffix was involved could have been subject to developments lying beyond what historians of Ossetian have found necessary to reconstruct up to this point. I would like to point to the existence of a set of nouns which have heretofore proven more or less recalcitrant etymologically, but which may benefit from the positing of *-č.

Several forms with stems ending in -s appear to be related on formal and semantic grounds to stems which may be reconstructed as ending in *-t- or *-d-. Since such an alternation of *t/d with *s (or *š) would be difficult to explain easily in terms of Old Iranian phonology and morphology, the existence of this relation, let alone its significance, has not been noted hitherto. Nevertheless, the suffix *-č which we have posited opens the door to reading the s of these forms as a development arising from a stem-final dental in contact with the suffix.

Miller (1907: 332-333) took Ossetian *us/(w)osæ* 'woman, wife' to be an Iranian equivalent of Vedic *yosā-* (or *yosit-*) 'young woman.' While this claim is plausible phonologically, it would be more compelling if there were evidence for the existence of this stem elsewhere in Iranian. I believe that it would be a mistake to attempt to separate the Ossetian form from the familiar Indo-Iranian stem **vadhū-* 'woman' (cf. Sanskrit *vadhū-*, Avestan *vaðū-*, Sogdian *wǎ(w)*, etc.). If we assume an umlaut like the one leading to the mid-vowel seen in *midæg/medæg* 'in, internal' (< **madyaka-*), a Pre-Ossetian form **wod-* < **vawd-* < **vadv-* becomes entirely conceivable. It is not beyond possibility that

the *-č- suffix could have merged with the final *ǝ of the stem to give rise to something that would show up as -s in modern Ossetian.

An example of a phonological phenomenon very similar to what we have envisioned in the case of *us* may be seen in *bos/bo(n)s* 'leash, tether.' Abaev tentatively interprets this form as a borrowing from Turkish *boş* 'lace,' but the presence of *n* in the Digor reflex is left unexplained in such an analysis. Given *us/osæ* < **vadu-*, it is quite possible that *bo(n)s* is to be traced back to the Iranian verbal root **band-* 'tie' (cf. modern Ossetian verb *bætt-yn/bætt-un*), with the same development of *-d-č- > -s as is seen in the word for 'woman.'

We may likewise explain the Digor form *ræwes*, *ræwves*, *ræwbes* 'roe-deer' by appealing to the same development of *-d-č- > -s. If we posit that the front vowel of the second syllable is due to the effect of the umlaut which have seen above (cf. the vowel of Digor *festæ* 'wool from the spring shearing' < **fans-č* above), it is quite possible to relate this form to *ræwvad* 'swift-footed,' a compound of *ræw* 'swift' and *fad* 'foot.' The addition of the suffix may be ascribed to the small size of the roe-deer compared with other types of deer.

A similar case may be found in a word from livestock terminology. Ossetian *fos/fons* 'cattle, wealth, loot,' which has the same general configuration as *bos/bo(n)s*, is clearly akin to Avar *panz* 'horned cattle,' Ingush *fos* 'loot,' and Chechen *hons* 'loot.' Abaev suggests, "... In the absence of a satisfactory Iranian etymology it remains to admit that the Ossetian word is a borrowing from Avar, and the Chechen and Ingush from Ossetian..." (1958: 479).

We have seen, however, that Iranian etymologies for Ossetian words are at times not obvious. Taking the final -s of *fo(n)s* to be the result of the co-occurrence of a stem-final dental and the *-č-suffix, we find that we have the opportunity to take this word to be based upon a Pre-Ossetian stem **fand-*. While such a form does not afford an immediate and unequivocal solution to the question of the history of *fo(n)s*, it does provide avenues for exploration. It might be asked, for example, whether **fand-* is to be related in some manner to Iranian **payah-* 'milk' (perhaps, e.g., ultimately < **paya(h)-vant-* 'milk-possessing'?). Alternatively, given the connection between 'cattle' and 'wealth' seen in the two senses of *fo(n)s*, we might choose to see in *fo(n)s* the outcome of a derived correlate to Ossetian *fænd* 'intention,

plan, purpose, desire.¹²

Whatever the ultimate source of **fand-* may have been, it is possible that we find a trace of it (in the vrddhi-grade **fānd-*) in Ossetian *nyvond* 'sacrifice,' *nyvondag* 'animal destined for sacrifice' (which Abaev (1973: 215) derives from the verbal stem *band-* 'tie'). Indeed, if we relate **fānd-* to Ossetian *fænd* it is not difficult to imagine the circumstances under which 'desire' could have come to mean 'animal' in the context of religious ritual: the verbal expression *nyvond kænyn* lit. 'to do a *nyvond*' (< **ni-fānd- kan-un*) could have developed from an original meaning 'to express a desire' to 'to undertake a sacrifice' by a fairly simple reinterpretation, with a simple metonymy leading from **fand* 'desire' > 'sacrificial beast' > 'cattle.'

It is possible that we find an instance of a diminutive suffix attached to a labial-final stem in the pair *kæf* '(large) fish' vs. *kæsaɡ* (Digor *kæsalgæ*) 'fish.' The former term is known to be of Iranian origin (cf. Sogdian *kp*, Khotanese *kava-*, Pashto *kab*, Yidgha *kap*, Munjan *kâp*, Wakhan *kup*), but *kæsaɡ/kæsalgæ* remains etymologically unclear. The *š* of Georgian *kašaqi* 'herring' suggests that the sibilant of *kæsaɡ* was previously something other than simple *s* (Bielmeier 1977: 167-169). If we posit an earlier **kaf-č-* '(small) fish,' with an outcome of the sequence **-f-č-* comparable to what we have seen among the dental-final stems, we could have an explanation of *kæs-*.¹³ If this is indeed the source of *kæsaɡ/kæsalgæ*, the developments posited must have happened at a quite early point in history, if Hungarian *keszeg*, Ostyak *kašeu* and Vogul *käsen* 'bream' were borrowed from Alanic with a simple sibilant.¹⁴

¹²Iranian **pant-* (Ossetian *fænd*, Persian *pand* 'opinion, advice') is taken by Abaev to be a semantic development of **pant(h)ah-* 'road' (1958: 445). I suspect that it might better be viewed as a cognate of Sanskrit *panita-* 'admired,' the participle of *pan-*, with the loss of medial schwa normal for Iranian.

¹³If Georgian *kašaqi* evidences a trace of a sibilant distinct from normal Ossetian *s* (i.e., a sibilant ultimately from **-f-č-*), the problem might in some manner be related to the unexpected *š* of the Xevsurian personal name *Apšina* < Ossetian *æfsin(æ)*, which similarly lies in the vicinity of an *f* (Bielmeier 1977: 168).

¹⁴Is it possible that Ossetian *kajys*, *kais* (Digor *kajes*) 'father-in-law of the wife (Russ. *test'*) and the members of his family' also contains a final *-s* which is somehow to be traced back to an expressive **-č?* Abaev's explanation of this word as coming from **ka + es* 'he who is' (on the grounds of a taboo proscribing the mentioning of the name of the father-in-law) is not satisfying. The similarity between the first syllable of the Ossetian form and that of the corresponding

There are, in sum, a number of individual lexical items in the modern Ossetian lexicon which may potentially be analyzed on the basis of a reconstructed suffix **-č-*. As we have seen, in a fairly large number of cases phonological developments have intervened to obscure the original relationship between basic and suffixed stems. There remains no doubt a fair number of nominal and adjectival forms yet to be identified which could benefit from investigation along these and similar lines, but we leave them aside at this point to turn briefly to address the question of the meager data available on the language of the Scythians and Alans, the linguistic precursors of the Ossetians. Although these peoples played extremely important roles in the history of the Eurasian steppe, they left no written records of any great value for the historian of the grammar of their language. We are therefore obliged to rely upon the scant information left by the cultures bordering upon the steppe, in particular the Greek and the Muslim. In certain of the linguistic details left by these secondary sources, I would like to suggest that we find a measure of confirmation for our conjectures concerning a suffix **-č-*.

The twelfth-century Greek poet Johannes Tzetzes incorporated into his "Theogony" a number of phrases in the languages of the foreign peoples familiar to the Byzantine court. Among these phrases he included two lines of Alanic, together with Greek glosses (Moravcsik 1930; Hunger 1953).

*... toîs alanoîs prospthéggomai katà tèn toútōn glōssan
kalē hēméra sou authénta mou arkhóntissa póthen eîsai
tapagkhàs méspili khsinà korthi kantà; kai tálla
án d' ékhē alánissa papân phílon akoúsais taúta
ouk aiskhúnesai authéntria mou nà gamē tò mounín sou papàs
to phárnetz(nlē) kintzi méspili kaitz phouà saoûgge*

... I greet the Alans in their language.
Good day to you, my lord, lady, where are you from?
taparxas mesfili xsina korthi kanta and other things.
When an Alan woman has a priest as a friend you hear this:

Turkic forms (e.g., Turkish *kayn-ata*) is conspicuous. Could early Ossetian **kajes* have developed in some manner from a **-č-*-suffixed form of a Turkic loanword (i.e., Turkic **qayin-ata-* > Oss. **kayin-at(a)-č* > **kaj(n)-es*?)? Cf. the suffixed Turkic loanword *čyzg/kizgæ* 'daughter.'

Are you not ashamed, my mistress, that a priest has relations with you?
to farnetz(n/i) kintzi mesfili kaitz fwa saw(ŋ)ge...

It is possible to identify much of Tzetzes' Alanic on the basis of modern Ossetian—most notably, *tapagkhas*, glossed as 'Good day to you,' corresponds neatly to the modern Ossetian greeting *dæ bon xorz* ('Your day (be) good'). The short text which Tzetzes provides happens to contain two of the words which have already figured in our discussion of the suffix *-č-, *čynz/kinzæ* (= Tzetzes' *kintzi*) and *æxsin(æ)* (= *khsina*); the forms are virtually identical to the Digor shapes of the modern words.

A more intriguing detail of even greater relevance for the issue of *-č- may be found in the second line of Alanic, in the words *to phárnetz(n/ē)*. Morávcsik observes that the final letter of *pharnetz* might be either an *nu* or an *ēta*, the latter of which in the twelfth century would presumably have been pronounced as *i*; in Hunger's manuscript this character is absent. Abaev took the stem of the form to be the same as in Ossetian *farn* 'glory, splendor,' but he was unable to say anything specific about *pharnetz(n/ē)*. Taking it to correspond in some manner to the Greek 'you are not ashamed,' he pointed out that "...The indubitable semantic connection between the word *farn* and the feeling of shame, which emerges in this use, is also preserved, evidently, in the expression presented by Tzetzes, if the translation given by him is correct" (Abaev 1949: 259).

Bielmeier (1993: 16-17), on the other hand, emends *phar-* to **sfar(m)* (or **fsar(m)*), to be compared to Iron *(æ)fsarm*, Digor *(æ)fsar(m)*, Buddhist Sogdian *šβ'r*, Manichaean Sogdian *šf'r*, Khwarezmian *spr̥m*, Yidgha *fšarm/šfarm*, Late Avestan *fšarəma-*, all of which mean 'shame.' The following characters *netz(n/ē)* he interprets as a counterpart either to Ossetian *nicy/neci* 'nothing' or to Ossetian *næj* 'is not' + *zy/zi* 'therein'. In either case, the sense of the passage would be on the order of 'ist nicht Scham darin?' This is not implausible, but it requires, in addition to the emendation of *ph-* to **sph/*phs-*, the assumption that the loss of the final *-m* of Digor *(æ)fsar(m)* had taken place already by the twelfth century.¹⁵

If our speculations about the existence of an early feminine suffix

¹⁵The interpretation of *tz(n/ē)* as *zi* is improbable, since one would expect to find the clitic pronoun in Wackernagel's position directly following *to*.

are correct, we would expect to find that at some earlier stage the pre-Ossetian adjective agreed in gender with its noun. The ending *-etz* of *pharnetz* (assuming that Hunger's text presents the original shape of the word) is a plausible shape for the feminine suffix **-č* to have taken in twelfth-century Alanic. It is therefore possible that we have here a feminine counterpart to the ancestor of modern Ossetian *farnyg* (Digor *farnug*) 'possessing glory,' serving as the predicate to the second-person pronoun *to* (= modern Ossetian *dy*, Digor *du*). By this reading the Alanic phrase lying behind Tzetzes' 'are you not ashamed?' would therefore have been a rhetorical question *du farniz* 'are you honorable (f.)?'

Another problem for interpretation lies in the following words, *kintzi mesphili*. The first word *kintzi* is clearly related to Ossetian *čynz/kinzæ* 'bride,' and Abaev (1949: 258) assumes that *kintzi mesphili*, which apparently corresponds to 'my mistress' in the Greek gloss, is some form of apposition ("wife—my lady"), despite the fact that *mesphili* in the preceding line of Alanic is clearly glossed as masculine 'lord.' Bielmeier (1993: 6-13) prefers to take *mesphili* to be a compound, the second element of which is a reflex of Iranian **parvya-* 'first.'¹⁶ The entire sequence *kintzi mesphili*, in Bielmeier's interpretation, is to be read as 'die junge Frau, die Herr(in) ist'; the construction "head noun (genitive) + modifier" is still found to some degree in modern Ossetian, as in *mad-y zæronð* 'old mother' alongside the more frequent *zæronð mad*.

An alternative interpretation of *kintzi mesphili* is available if we assume that in the Alanic of Tzetzes' time genitives could still follow their heads. We may thus read the sequence *kintzi mesfili* as the Alanic counterpart to 'my mistress' by taking it to be *kinzi me-sphili*, literally 'bride of my *sphili*,' where *sphili* has the meaning 'lord' seen in the preceding line of Alanic. Given the presence in modern Ossetian of the *ezāfe*-like construction *mad-y zæronð* (see above) with a modifying adjective, it is not unlikely that a genitival *ezāfe* could have

¹⁶Bielmeier explores a number of options in the interpretation of *mes-*, but seems to settle on a derivation from **maiθa-* 'house,' with a development of **θ >* sibilant as in eastern Yaghnobi (rather than to *t*, as in modern Ossetian). *Mes-phili* would therefore be read as "erster des Hauses" (1993: 13). It seems preferable, however, to see in *me-* some reflection of the "my" of the gloss.

existed in the older stage of the language (**kinz-i me-sfil-i*).¹⁷

It is possible that we find an example of considerably greater antiquity in Herodotus' fourth book. Among the major rivers of Scythia (i.e., the southern Russian steppe region) Herodotus mentions the *Pantikapēs*.

... East of the Borysthenes [= the Dnepr], starting from the sea one comes first to Hylaea—the Woodland—to the northward of which are Scythians who get their living from the land and are known to the Greeks on the river Hypanis [= the Bug] as Borysthenites. They call themselves Olbiopolites. These Scythians extend eastward as far as a river named the Panticapes—the distance of a three days' journey... Eastward of the Scythians who lived off the land, and on the other side of the Panticapes, are the nomadic Scythians, who know nothing of agriculture... (IV: 18-19)

The fifth river, the Panticapes, also rises in a lake and flows in a southerly direction through Hylaea—the Woodland—to join the Borysthenes. The country between it and the Borysthenes is occupied by the Scythian agricultural tribes... (IV: 54)

The name *Pantikapēs* has been taken to be a Scythian compound consisting of a reflex of Indo-Iranian **pant(h)-* 'road' (cf. Ossetian *fændæg* 'road') and **kapa-* 'fish' (cf. Ossetian *kæf* 'fish') (Abaev 1949: 175). While this possibility certainly cannot be ruled out, we should also observe that the break between the two components of the compound may also be drawn differently. If we allow for the possibility that the syllable *-ka-* of *Pantikapēs* represents the well-known Iranian suffix, we see that the remainder of the word is distinctly reminiscent

¹⁷In the remainder of the line, *kaitz phouà saouŋge*, corresponding to which the Greek has *nà gamē tò mounin sou papàs*, Bielseimer sees (1) a verbal stem **qaj-* 'future' (to be compared to modern Ossetian *qæj-yn*, Khwarezmian *γγ-*, Wakhi *γγ-*) provided with the verbal noun formant *-c*, (2) an early counterpart to the subjunctive verb 'be' (= modern Ossetian *fæ-wa*), (3) a counterpart to modern Ossetian *sawŋyn* (Digor *sawgin*) 'priest.' The presumed syntax of the clause under this interpretation remains unclear to me. I would prefer to interpret *kaitz* as an archaic correspondent to the modern relative pronoun *kæcy/kæci* < **ka-* 'who' + **či-*; in Tzetzēs' form *kaitz* we seem to have a counterpart to *kæj* (Iron) or *ke* (Digor), the genitive of *čilka* 'who,' + the particle *-c(i)*—i.e., in Tzetzēs' form the declension still resides in the pronoun (gen. *kaj-c(i)*) rather than in the entire complex, as it does in modern Ossetian (*kæcy-jy*, gen. of *kæcy*). The whole line *to pharnetz kintzi mesphili kaitz phoua saouŋge* can therefore be read as *du farniz, kinzi me-sfili, kajc(i) fawā sāwgi(n)?* 'are you honorable, O bride of my lord, who will have ("whose is to be") a priest?'

of Indo-Iranian **ap-* 'water, river.' The name of Herodotus' fifth river could thus quite easily be read as **Pantika-ap-*. If this reading is correct, in the ending *-ika* of the first element we may well have an attestation of the forerunner of the feminine suffix which we have been reconstructing. Indeed, although it may very easily be coincidence, the form *pantika-* is strikingly reminiscent of the shape we would choose to reconstruct the original source of Ossetian *fo(n)s* 'cattle.'

The identity of Herodotus' *Pantikapēs* has long been disputed, the prevailing view being that Herodotus had the modern Ingulec in mind. Space does not allow us to delve into the arguments surrounding Herodotus' Scythian geography at length, but it is worth noting that Tomaschek has proposed taking the *Pantikapēs* to be the Samara (Tomaschek 1893): like the *Pantikapēs* (and unlike the Ingulec), the Samara lies to the east of the Dnepr and enters the latter from the north, even though it has its source in the south (in what Herodotus called *Hylaea*). Once again, while it may be a matter of coincidence, the westernmost tributary of the Samara is still called *Pansova*, a form which makes a strikingly plausible reflex of **Pantika-ap-*, given our speculations about the interaction of the suffix **-č-* and stem-final dentals: **pant-ika* - > **pant-(i)ča-* > **pans-*, while **-aβ* < **ap-* might easily result in *-ov-* in a Slavic-language environment.¹⁸

If *Pansova* does indeed come from *Pantikap-*, at least certain of the developments which we have discussed here must have been extremely early—the change **-t-* + **-(ⁱ)č-* > *-s*, in any event, must presumably have occurred while some form of Scythian was still spoken in the lower Dnepr basin, i.e., by the third century A.D. On the other hand, if our reading of Johannes Tzetzes is correct, as late as the twelfth century pre-Ossetian could still mark gender agreement in adjectives. In the preceding pages we have suggested that at one point the ancestor of Ossetian used a form of suffixation to mark the feminine gender, in a manner not greatly different from what other Middle Iranian languages, such as Sogdian and Khwarezmian, were doing at the same time, and that this suffix was also employed as an expressive derivational marker. This system of gender-marking through suffixation is now as dead as the Old Iranian system which preceded it, but it has left traces of itself deeply embedded in the lexicon of the modern language.

¹⁸Trubačev (1968: 261) interprets the name *Pansova* as < Iranian **pansava-* 'sandy' (cf. Avestan *pąsu-* 'dust')

REFERENCES

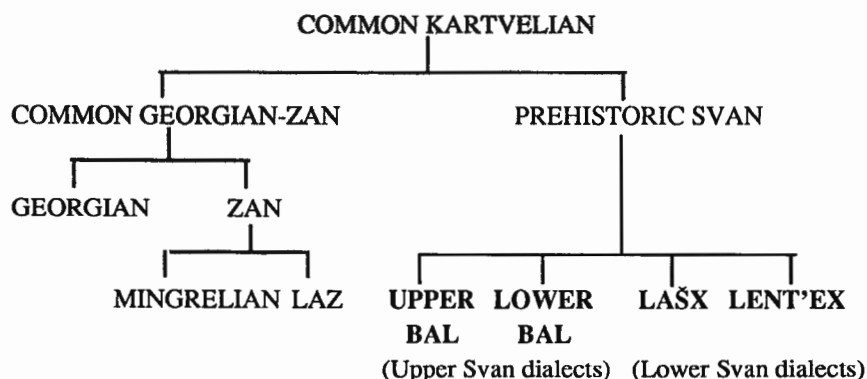
- Abaev, V. I. 1949. *Osetinskij jazyk i fol'klor*. Moscow-Leningrad: AN SSSR.
- Abaev, V. I. 1958. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka. I: A-K'*. Moscow-Leningrad: AN SSSR.
- _____. 1973. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka. II: L-R*. Leningrad: "Nauka."
- _____. 1979. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka. III: S-T'*. Leningrad: "Nauka."
- _____. 1989. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka. IV: U-Z*. Leningrad: "Nauka."
- Bartholomae, Christian. 1961. *Altiranisches Wörterbuch* (2. unveränderte Auflage). Berlin: Walter de Gruyter & Co.
- Benveniste, E. 1959. *Études sur la langue ossète*. Paris: C. Klincksieck.
- Bielmeier, Roland. 1977. *Historische Untersuchung zum Erb- und Lehnwortschatzanteil im ossetischen Grundwortschatz*. Frankfurt am Main/Bern/Las Vegas: Peter Lang.
- _____. 1993. "Das Alanische bei Tzetzes," in Wojciech Skalmowski and Alois van Tongerloo, edd. *Medioiranica: Proceedings of the International Colloquium organized by the Katholieke Universiteit Leuven from the 21st to the 23rd of May 1990 (Orientalia Lovaniensia Analecta 48)*. Leuven: Peeters. pp. 1-28.
- Gershevitch, Ilya. 1954. *A Grammar of Manichean Sogdian*. Oxford: Basil Blackwell.
- _____. (ed. Nicholas Sims-Williams). 1985. *Philologia Iranica*. Wiesbaden: Ludwig Reichert.
- Herodotus (ed. Reginald Walter Macan). 1895. *Herodotus: The Fourth, Fifth, and Sixth Books*. London: Macmillan.
- Humbach, Helmut. 1989. "Choresmian," in *Compendium Linguarum Iranicarum* (ed. Rüdiger Schmitt). Wiesbaden: Ludwig Reichert.
- Hunger, H. 1953. "Zum Epilog der Theogonie des Johannes Tzetzes." *Byzantinische Forschungen* 46: 302-307.
- Miller, Wsewolod. 1903. *Die Sprache der Osseten (Grundriss der iranischen Philologie: Anhang zum ersten Band)*. Strassburg: Karl J. Trübner.
- Miller, Wsewolod. 1907. "Beiträge zur ossetischen Etymologie." *Indogermanische Forschungen* 21: 323-334.
- Moravcsik, J. 1930. "Barbarische Sprachreste in der Theogonie des Johannes Tzetzes." *Byzantinisch-neugriechische Jahrbücher* VII/3-4: 352-365.
- Reichert, Hans. 1978. *Awestisches Elementarbuch* (dritte, unveränderte Auflage). Heidelberg: Carl Winter.
- Tomaschek, Wilhelm. 1893. "Akesinos," in *Paulys Real-Encyclopädie der classischen Altertumswissenschaft* (Georg Wissowa, ed.). Stuttgart: J. B. Metzlerscher Verlag.

Trubačev, O. N. 1968. *Nazvanija rek pravoberežnoj Ukraïny*. Moscow: Izdatel'stvo "Nauka."

AORIST AND PSEUDO-AORIST FOR SVAN ATELIC VERBS

K. Tuite
Université de Montréal

0. Introduction. The Kartvelian or South Caucasian family comprises three languages: Georgian, which has been attested in documents going back to the 5th century, Zan (or Laz-Mingrelian) and Svan. In the case of Zan and Svan, there is almost no documentation going back beyond the mid-19th century. There is a consensus among experts that Svan is the outlying member of the family, whose family tree would be as in this diagram:



In this paper I will look at one segment of the formal system in Svan — in particular, in the more conservative Upper Svan dialects— for coding aspect. In this introductory section I will go over the system of what Kartvelologists term ‘screeves’ (sets of verb forms differing only in person and number), and then the classification of verbs by lexical aspect with which the morphological facts will be compared.

0.1. Screeves and series. In the morphology of verbs in each of the Kartvelian languages, one of the fundamental distinctions is between verb forms derived from the SERIES I OR ‘PRESENT-SERIES’ stem, and forms derived from the SERIES II OR ‘AORIST-SERIES’ stem. We will not go into all of the formal differences between these two stems for the various groups of verbs in each language, save to note that for most verbs the Series I stem contains a suffix (the ‘series marker’ or ‘present/future stem formant’) which does not appear in Series II forms. (There is also a third series of verb forms, the ‘perfect series,’ in the Kartvelian

languages. These forms appear to be of relatively recent origin, and in any event are built on recognizably Series I or II stems. Series III will not be discussed in this paper.)

In Old Georgian (5th-11th centuries) those screeves employing Series II stems were aspectually durative or *linear* (Geo. *xazovani*), while those verb forms built on Series I stems were characterized by *punctiliar* (Geo. *c'ert'ilebrivi*) aspect. That is, the system offered a choice in the representation of an event: it could be viewed as extending over a period of time, or the narrative spotlight could be directed, in a sense, at a salient point, a change of state. The nature of the semantic distinction conveyed by the formal opposition of Series I and II screeves in the modern Kartvelian languages has changed somewhat from that of Old Georgian (though relics of the older aspectual system are preserved in all branches of the family [Mač'avariani 1974]).

The screeves of Svan and, for the purpose of comparison, Modern Georgian, derived from Series I and Series II stems are shown in {1} [Gudjedjani and Palmaitis 1986] (the Zan languages will not be discussed here):

{1} MODERN GEORGIAN AND SVAN SCREEVES

Series I ('present series')

Syntactic: 1st and 3rd conjugation verbs assign NOM case to their subjects

Morphological: Special suffix ('series marker') present in Series I stems

| Modern Georgian | | | Svan | | | |
|-----------------|---------------------|------------------------|------------------|-------------|--------------|---------------------|
| (1) | (2) | | (1) | (2) | (3) | (4) |
| -PST | present | future | present | prfv future | impfv future | narrative present |
| +PST | imperfect | conditional | imperfect | prfv condl | impfv condl | narrative imperfect |
| MODAL | present conjunctive | con-future conjunctive | prs con-junctive | | | narr. con-junctive |

Svan examples: [Upper Bal; Gudjedjani & Palmaitis 1986]:

Series I stem *g=em* 'build'

pres. *a-g=em* 's/he builds it'

imperfective fut. *a-g=m-un-i* \Leftarrow ¹ *a-g-em-un-i*

¹Svan grammar has a lively morphophonemic component. Where clarity of

imperfect *a-g=äṁ-da* \Leftarrow ?*a-gem-w-da*
 perfective future *ad-g=em-n-e* \Leftarrow *ad-a-g-em-in-e*

Series II ('aorist series')

Syntactic: 1st and 3rd conjugation verbs assign ERG case to their subjects

Morphological: Series marker absent in Series II stems; sometimes distinguished by ablaut

| | <u>Modern Georgian</u> | | <u>Svan</u> |
|-------|------------------------|---------------------|-------------|
| | (1) | (2) | |
| +Pst | imperfective aorist | perfective aorist | aorist |
| Modal | imperfective optative | perfective optative | optative |

Svan examples [Upper Bal dialect; Gudjedjiani & Palmaitis 1986]:
 Series II stem *g* 'build'

aorist *ad-g-e* \Leftarrow *ad-a-g-e* 's/he built it'

optative *ad-g-a-s* \Leftarrow *ad-a-g-a-s*

0.2. Conjugation classes. Svan verbs can be divided into four conjugation classes based on certain morphological and syntactic criteria. The system employed here is based upon that devised by Topuria [1967:39-41], with some modifications adapted from Harris [1985]. The 1st and 2nd conjugation classes are largely composed of paired transitive and intransitive verbs, distinguished (usually) by affixation or ablaut. The Series II forms of 1st conjugation verbs assign ergative case to the NP controlling subjective agreement, while 2nd conjugation verbs assign nominative case.

{2}

1ST CONJUGATION

ä-č'm-e 's/he mows hay'

a-hrăq'-i 's/he brews vodka'

i-šx-i 's/he burns his/her own sthg'

pxiž-e 's/he spreads sthg'

kwic-e 's/he cuts sthg'

2ND CONJUGATION

i-č'm-i '[hay] is mowed'

i-hrăq'-i '[vodka] is brewed'

i-šx-i 'sthg burns'

pxež-n-i 'sthg is spread, scattered'

kwec-n-i 'sthg is cut'

3rd and 4th conjugation verbs are almost always intransitive, and do not come in pairs. Many 3rd conjugation verbs have stems containing

presentation demands it, surface forms are followed by representations of the underlying morphology, signalled by \Leftarrow .

the frequentative/durative suffixes *-ä:l-* or *-ie:l-*. In Georgian, the Series II forms of 3rd conjugation verbs assign ergative case to their subjects; the syntax of these verbs in the Svan dialects shows more variety, as we shall see. Here are some examples:

{3}

3RD CONJUGATION

sip' 'sb/sthg turns'

q'u:li '[cow] moos'

i-yr-ä:l 'sb sings'

i-gi:c'-ä:l 'sb/sthg swings'

i-q'wi:l-ie:l '[goat] bleats'

4TH CONJUGATION

sk'ur 'sb is seated'

tera 'sb/sthg is visible'

x-a-c'əx 'sb needs sthg'

x-o-šgur 'sb is ashamed'

x-o-xal 'sb knows sthg'

0.3. Telic and atelic verbs. The verb-stem morphology of the Kartvelian languages reflects, fairly closely, the lexical-semantic distinction between TELIC and ATELIC VERBS [Holisky 1981; Harris 1985]. The 1st and 2nd conjugation classes of Svan and Georgian are composed largely of telic verbs, that is, verbs whose meanings include a *salient change of state*. The Vendler classification scheme employed by most semanticists recognizes two varieties of telic verbs, termed *accomplishments* (e.g., the English verbs 'teach,' 'break_{tr},' 'kill,' 'give') and *achievements* (e.g., 'learn,' 'break_{intr},' 'die,' 'arrive,' 'stand up').

The 3rd and 4th conjugation classes contain (again, for the most part) verbs of ATELIC aspect. Atelics are opposed to telic verbs in that their meanings lack a salient change of state. There are two types of atelic verbs: *statives* (e.g., 'have,' 'know,' 'be broken,' 'be red,' 'be standing') and *activity verbs* (e.g., 'sing,' 'gallop,' 'glow,' 'misbehave').

In this paper I will look at some of the morphological characteristics of the Series II verb forms in Svan. I will attempt to show that in Prehistoric Svan atelic verbs lacked Series II screeves. At a later point in the prehistory of Svan these verbs acquired a 'pseudo-aorist,' a Series I screeve derived from the imperfect by addition of a preverb. In the most recent stage, the pseudo-aorist forms of atelic verbs took some on the syntactic and morphological characteristics of the (true) aorists of telic verbs. This process has taken somewhat different directions in the modern Svan dialects.

§1. The morphology of Series II in Svan. Most Svan aorists are conjugated in one of two ways, which I will term STRONG (athematic,

with ablaut) and WEAK (thematic, without ablaut) inflections. Both inflection patterns employ two distinct stems, one for the 1st and 2nd singular subject ($S_{1/2sg}$) form, and one for the 3rd singular subject and all persons in the plural ($S_{3/pl}$), as in the following Upper Bal paradigms [Topuria 1967:137,141; the Prehistoric Svan forms, marked with a single or double asterisk, are from Kaldani 1978]:

| | |
|---------------------|--|
| {4} | WEAK AORIST (1ST CONJUGATION VERB) |
| S_{1sg}/O_3^2 | $\dot{z}o=xw-\dot{z}wem \leftarrow \dot{z}i-an=xw-o-\dot{z}öm < *- \dot{z}om-i$ 'I wrecked sb's sth' |
| S_{2sg}/O_3 | $\dot{z}o=xw-\dot{z}wem \leftarrow \dot{z}i-an=x-o-\dot{z}öm$ |
| S_{3sg}/O_3 | $\dot{z}o=xw-\dot{z}om-e \leftarrow \dot{z}i-an=x-o-\dot{z}om-e < *- \dot{z}om-i-a$ |
| $S_{1pl, incl}/O_3$ | $\dot{z}o=lw-\dot{z}om-e-d \leftarrow \dot{z}i-an=l-o-\dot{z}om-e-d$ |
| $S_{1pl, excl}/O_3$ | $\dot{z}o=xw-\dot{z}om-e-d \leftarrow \dot{z}i-an=xw-o-\dot{z}om-e-d$ |
| S_{2pl}/O_3 | $\dot{z}o=xw-\dot{z}om-e-d \leftarrow \dot{z}i-an=x-o-\dot{z}om-e-d$ |
| S_{3pl}/O_3 | $\dot{z}o=xw-\dot{z}om-e-x \leftarrow \dot{z}i-an=x-o-\dot{z}om-e-x$ |
| {5} | STRONG AORIST (1ST CONJUGATION VERB) |
| S_{1sg} | $o=\check{c}'k'or \leftarrow a=xw-\check{c}'k'or < *- \check{c}'k'or-\emptyset < **-\check{c}'k'or-i$ 'I cut sthg' |
| S_{2sg} | $a=\check{c}'k'or \leftarrow a=x-\check{c}'k'or$ |
| S_{3sg} | $a=\check{c}'k'wer \leftarrow a=\check{c}'k'ör < *- \check{c}'k'or-i < **-\check{c}'k'or-i-a$ |
| $S_{1pl, incl}$ | $a=l-\check{c}'k'wer-d \leftarrow a=l-\check{c}'k'ör-d$ |
| $S_{1pl, excl}$ | $o=\check{c}'k'wer-d \leftarrow a=xw-\check{c}'k'ör-d$ |
| S_{2pl} | $a=\check{c}'k'wer-d \leftarrow a=x-\check{c}'k'ör-d$ |
| S_{3pl} | $a=\check{c}'k'wer-x \leftarrow a=\check{c}'k'ör-x$ |

There is no stem alternation in the Svan optative screeve; for most verbs the optative stem is identical to the $S_{1/2sg}$ aorist stem.

According to the reconstruction proposed by Kaldani [1978], both inflection patterns, strong and weak, derive from the same Prehistoric Svan paradigms. For transitive verbs (almost all of which have 1st conjugation morphology in the modern dialects), the aorist desinences were $*-i$ in the $S_{1/2sg}$ stem, and $*-i-a$ in the $S_{3/pl}$ stem ($*-a$ is descended from the Common Kartvelian S_{3sg} past-indicative ending, later extended

²In glossing the person marking of Kartvelian verbs, I employ *S* for "subject" and *O* for "object" affixes, with subscripts indicating person, number and the inclusive/exclusive distinction.

to all plurals in the aorist). The aorists of intransitive verbs (i.e., 2nd/3rd/4th conjugation verbs in the modern dialects) lacked the **-i* suffix; their stem formants were **-Ø* in the $S_{1/2sg}$ stem, and **-Ø-a* in the $S_{3/pl}$ stem. The present-day strong aorists are descended from verbs which underwent loss of final vowels at a very early date, predating umlaut (marked with a double asterisk in the above paradigms). Weak aorists lost their final vowels somewhat later, after it had become a productive rule of Svan morphophonemics that deleted /i/ or /e/ triggered umlaut in the preceding vowel (the single-asterisk stage); hence the umlauted root vowels in the $S_{1/2sg}$ stems of weak aorists, and in the $S_{3/pl}$ stems of strong aorists.

1.1. Preverbs and aspect. The initial morpheme(s) in the verb forms in {4} and {5}, separated from the rest of the word by a double hyphen, are called **PREVERBS**. There are two sets of these prefixes in Svan: **OUTER PREVERBS** (e.g., *ži* 'up, upon' in {4}), which are comparatively loosely connected to the following morphemes, and **INNER PREVERBS** (e.g., *an* 'hither' in {4}, *a(d)* 'away' in {5}), which are morphophonemically integrated into the verbal complex. Although preverbs may add a directional meaning to the significance of the verb, in many cases their semantic contribution is no more transparent than that made by their counterparts in the Indo-European languages (e.g., Russian *so=znat* 'realize' [cf. *s(o)* 'with,' *znat* 'know']; English 'put up (with),' 'take over').

In Svan, as in Modern Georgian—and similarly to the Slavic languages—preverbs play a role in the signalling of aspectual distinctions. Svan and Modern Georgian 1st and 2nd conjugation verbs have paired Series I screeves, opposed (primarily) in perfectivity. The screeves in column (1) of fig {1} are imperfective, and those in column (2) are perfective. For most verbs the column (1) screeves lack preverbs, and the column (2) screeves have them, e.g.:

{6} **The aspect-marking role of preverbs in Series I**

Modern Georgian

| | IMPERFECTIVE ASPECT | PERFECTIVE ASPECT |
|--------------|-----------------------|--------------------------------|
| - <i>Pst</i> | present: c'er-s | future: da =c'er-s |
| + <i>Pst</i> | imperfect: c'er-d-a | cond/iter: da =c'er-d-a |
| | 's/he was writing it' | |

Svan (Upper Bal)

IMPERFECTIVE ASPECT PERFECTIVE ASPECT

| | | | | |
|--------------|------------|-------------------------|--------------|-----------------------------|
| - <i>Pst</i> | present: | t'ix-e | prfv future: | ä=t'xe ← an=t'ix-e |
| + <i>Pst</i> | imperfect: | t'ix-a | prfv condl: | ä=t'xa ← an=t'ix-a |
| | | 's/he was returning it' | | |
| - <i>Pst</i> | present: | a-gem | prfv future: | ad=gem-n-e ← ad=gem-in-e |
| + <i>Pst</i> | imperfect: | a-gäm-da | prfv condl: | ad=gem-n-a ← ad=gem-in-a |
| | | 's/he was building it' | | |

In the case of most Modern Georgian 1st and 2nd conjugation verbs, and many of their Svan counterparts, the presence or absence of a preverb is the *sole* indicator of aspect. For many other Svan 1st and 2nd conjugation verbs, perfective Series I screeves contain an additional suffix as well as a preverb [Topuria 1967:115-130; Čumburidze 1986: 184-215].

In the formal marking of aspectual distinctions in Series II screeves, there is a significant difference between Svan, on the one hand, and Modern Georgian, Laz, and Mingrelian, on the other. The latter languages allow the same opposition of preverbed and preverbless forms in the aorist and optative screeves as is found in Series I (see {7}).

{7} The aspect-marking role of preverbs in Series II

Modern Georgian

IMPERFECTIVE ASPECT PERFECTIVE ASPECT

| | | | | |
|--------------|----------------|------------------------------|--------------|-------------|
| + <i>Pst</i> | impfv aorist | c'er-a | pfv aorist | da=c'er-a |
| <i>Modal</i> | impfv optative | c'er-o-s | pfv optative | da=c'er-o-s |
| | | 's/he would/should write it' | | |

The presence of a preverb indicates that the narrative focus is upon the *accomplishment* or *achievement* of the denoted action, while the absence of a preverb indicates that the focus is upon the period *preceding* accomplishment (whether or not this ever occurs). Imperfective aorists also can imply iteration in some contexts [Mač'avariani 1974:120-1]. Šanidze [1953:272] illustrated this distinction by citing some Georgian

proverbs which exploit the semantic difference between imperfective and perfective aorists, e.g., :

- {8} katamma čxrik'a, čxrik'a da tavisi dasak'lavi dana
gamo=čxrik'a-o.

'The chicken *pecked and pecked* {IMPFV AOR} and (finally)
pecked out {PFV AOR} the knife that will be used to kill it.'

For verbs having this opposition, the non-preverbed aorist (e.g., čxrik'a) specifically draws attention to the period when the expected result was not (yet) achieved, and the preverbed aorist (e.g., gamo=-čxrik'a) signals achievement (in this case, the chicken's pecking exposed the knife). Mingrelian imperfective aorists are used in similar contexts to the above (see the proverbs and riddles cited in Qipšidze 1914:183, #8-9), and also in conjunction with negative particles, as in this sentence from Saxok'ia [1988:268]:

- {9} č'k'ver osurk kot ko=kiminu, vart mitins orčkinu; xangak vart
kiminu do ki'anas kišiot'q'vinu.

'The smart woman *did it* {PFV AOR}, but did not show it to anyone; the foolish one *did not do it* {IMPFV AOR}, and let everybody know about it.'

Imperfective aorists and optatives are in fact awkward or unacceptable for many Georgian 1st conjugation and certain classes of 2nd conjugation verbs (so-called suffixal and root intransitives), and overall are used less often than their perfective counterparts, but they are by no means rare [see Mač'avariani 1974:120-1; Met'reveli 1988]. By contrast, the formal opposition between preverbless and preverbed Series II screeves appears to be absent in Svan [Topuria 1967:139; Mač'avariani 1974:137].³ There is no exact Svan parallel to the semantic distinction

³I have encountered no examples in texts, including Davitiani's [1973] collection of Lower Bal Svan proverbs. Asked to translate the Georgian proverb in {8}, a Svan informant gave me a paraphrase with only one verb (a preverbed aorist). According to Mač'avariani [1974:137], "in Svan, unlike the other Kartvelian languages, imperfective (preverbless) aorists are very rare. V. Topuria [1967:139] considers the preverb to be an inseparable component of the Svan aorist." In a footnote on the same page, Mač'avariani cites what he claims is the *sole* example he was able to find of a Svan imperfective aorist, devised by a Lašx informant:

between perfective and imperfective aorists and optatives characteristic of Modern Georgian and the Zan languages.⁴

1.2. The formation of Series II screeves for atelic verbs. The preceding discussion of Svan morphology has dealt only with verbs of the 1st and 2nd conjugation classes. As noted earlier, these formal classes predominantly consist of telic verbs, the semantic representations of which contain a prominent transition or change of state. The remaining conjugation classes (3rd and 4th) are composed almost exclusively of atelic verbs. Svan 3rd conjugation verbs are mostly activity verbs, and the 4th conjugation is made up of stative verbs. We will look at statives first, then activity verbs.

According to Gudjedjiani & Palmaitis [1986:88-90] 'static verbs' generally only appear in imperfective (i.e., preverbless) screeves: the perfective future and conditional of Series I and both Series II screeves are lacking. The exceptions are stative verbs with pseudo-aorists and pseudo-optatives formed by *the addition of a preverb to the imperfect and present conjunctive stems*, respectively; these forms, therefore, are based on Series I, not Series II, stems [Topuria 1967:156-161,166; Mač'avariani 1974:138]. The difference in meaning between imperfect and pseudo-aorist—to judge from the Georgian glosses provided by Topuria—ranges from durative vs. momentaneous (*x-a-lo:n-(d-a)* // *at=lo:n*) to state vs. inchoative (*x-a-lt'-ən-d-a* // *la=xlat'-ən-d-a*). Here are some examples:

{10} Preverbed screeves of Svan stative verbs (Upper Bal dialect)

| imperfect | <i>pseudo-aorist</i> |
|--|--|
| (1) <i>x-a-lo:n-(d-a)</i> 'was offended [impf]' [c'q'inda] | <i>at=lo:n</i> \Leftarrow <i>ad=x-a-la-ən-(d-a)</i> 'was offended [prfv]' [<i>ec'q'ina</i>] |
| (2) <i>x-o-xal-d-a</i> 'knew' [<i>icoda</i>] | <i>lo=x-xal-d-a</i> \Leftarrow <i>la=x-o-xal-d-a</i> 'seemed/thought (sthg was the case)' |

xočid i xočid 'he was hitting him and hitting him' (Georgian *urt' q'a da urt' q'a*). Note, however, that Lašx is a Lower Svan dialect, in many respects more heavily influenced by Georgian than the relatively isolated Upper Svan dialects.

⁴Those Svan aorists which are attested without preverbs are aspectually perfective. They are believed to be relics, pointing to an earlier stage of Svan when preverbs did not have aspectual meaning [Topuria 1967:155-6; Machavariani 1974:137,139; Schmidt 1985].

- | | |
|-----------------------------|--|
| | [<i>egona</i>] |
| (3) x-a-lt'-ən-d-a | la=xlat'-ən-d-a ⇐ la=x-a-lat'-ən-d-a |
| 'loved' [<i>uq'varda</i>] | 'fell in love' [<i>šeuq'varda</i>] |
| present conjunctive | <i>pseudo-optative</i> |
| (1) x-a-lo:n-de:d-s | at=lo:n-de:d-s ⇐ ad=x-a-la-ən-de:d-s |
| (2) x-o-xal-de:d-s | lo=xw-xal-de:d-s ⇐ la=x-o-xal-de:d-s |
| (3) x-a-lt'-ən-de:d-s | la=xlat'-ən-de:d-s ⇐ la=x-a-lat'-ən-de:d-s |

This phenomenon is limited to Svan; there is nothing like it elsewhere in Kartvelian. Apparent parallels with Georgian (e.g., *uq'varda* 'sb loved sb/sthg' and *še=uq'varda* 'sb fell in love with sb/sthg') are nothing more than that. The first is formally the imperfect of a 4th conjugation stative verb (*u-q'var-s*), and the second the aorist of a derived 2nd conjugation inchoative (*še=u-q'var-d-eb-a*); their morphological similarity is largely coincidental [Topuria 1967:159].

Most Svan activity verbs (3rd conjugation), in most dialects, inflect like weak 1st conjugation verbs in Series II and—also like 1st conjugation verbs—assign ergative case to their morphological subjects (e.g., *eĭ-ne:m ädšdira:le* 's/he-ERG played'). Preverbs are always present in the aorist and optative of Svan activity verbs; in this respect their morphology differs sharply from that of Georgian activity verbs, e.g.:

[11] Aorist (3rd conjugation verb) — Upper Bal

- | | |
|-----------------------|---|
| S _{1sg} | ät=w-šdir-ä:l ⇐ ad=xw-i-šdir-ä:l 'I played' |
| S _{2sg} | ät=šdir-ä:l ⇐ ad=x-i-šdir-ä:l |
| S _{3sg} | äd=šdir-a:l-e ⇐ ad=i-šdir-a:l-e |
| S _{1pl,incl} | a=l-šdir-a:l-e-d |
| S _{1pl,excl} | ät=w-šdir-a:l-e-d |
| S _{2pl} | ät=šdir-a:l-e-d |
| S _{3pl} | äd=šdir-a:l-e-x |

(cf. Mod Geo equivalents: *v-i-tamaš-e*, *i-tamaš-e*, *i-tamaš-a* ... **without** preverbs)

The morphology and syntax of activity verbs is not uniform throughout the Svan-speaking territory, especially in the Lower Bal dialect region.

We note, first of all, that one and the same verb may form its Series II stem according to different patterns in different dialects. The following activity verb forms its aorist according to the 1st conjugation pattern in the Ecer subdialect of Lower Bal, while its cognates in the two Lower Svan dialects contain the suffix *-an/a:n-*, used to form the aorists of non-ablauting 2nd conjugation verbs. The syntax varies accordingly: the Ecer verb assigns ERG case, and its Lower Svan counterparts assign NOM [examples from Topuria 1967:237].

| {12} Ecer (Lower Bal) | Lentex | Lashx |
|-------------------------|----------------------|---------------------------|
| pres: i-t'bən-äl-x | i-t'bən-äl-x | i-t'bən-a:l-x |
| aorist: äd=t'əbn-al-e-x | čwäd=t'əbn-an-x | čwed=t'əbn-a:l-a:n-x |
| ← ad=i-t'əbn-al-e-x | ← ču-ad=i-t'əbn-an-x | ← ču-ad=i-t'əbn-a:l-a:n-x |
| | 'they spat' | |

In the Laxamulan subdialect of Lower Bal Svan, athematic aorists, *without* umlaut, have been described for 3rd conjugation verbs, in particular activity verbs with stems ending in the frequentitive suffix *-äl/əl-* [Kaldani 1978; Harris 1985:119-123]. Harris notes that athematic forms exist alongside, and are semantically equivalent to, thematic aorists formed from the same verbs. The most striking difference between the two sets of verb forms is in their syntax: the thematic aorists assign ERG case (e.g., *eĵ-nem äd=šdiral-e* 's/he-ERG played'), but their athematic equivalents assign NOM case (*eĵ-i äd=šdiral* 's/he-NOM played'). The two patterns are shown in {13} (forms elicited from A. Chkadua; cf. Kaldani [1978:152]).

{13} Aorist paradigms (3rd conj. verb)—Laxamulan (Lower Bal)

| | THEMATIC | ATHEMATIC |
|-----------------------|---------------------|----------------------------------|
| S _{1sg} | ot=šdir-al | ot=šdir-al ← ad=xw-i-šdir-al |
| S _{2sg} | ät=šdir-al | ät=šdir-al ← ad=x-i-šdir-al |
| S _{3sg} | äd=šdir-al-e | äd=šdir-al ← ad=i-šdir-al |
| S _{1pl,incl} | o=l-šdir-al-e-d | o=l-šdir-al-d ← ad=l-i-šdir-al-d |
| S _{1pl,excl} | ot=šdir-al-e-d | ot=šdir-al-d |
| S _{2pl} | ät=šdir-al-e-d | ät=šdir-al-d |
| S _{3pl} | äd=šdir-al-e-x | äd=šdir-al-x |
| | 'they [ERG] played' | 'they [NOM] played' |

Kaldani [1978] argues that the athematic pattern derives from the paradigm he reconstructed for Svan intransitive verbs (i.e., *-Ø in the $S_{1/2sg}$ stem, and *-Ø-*a* in the $S_{3/pl}$ stem, with later loss of the final vowel yielding an athematic $S_{3/pl}$ stem). Harris [1985:134-140] accepts Kaldani's hypothesis, and provides further arguments to support the view that the (presently) athematic paradigm for 3rd conjugation verbs is more ancient than the pattern with $S_{3/pl}$ -*e*). The presence of this latter inflectional paradigm in Laxamulan, and, as the dominant pattern for 3rd conjugation verbs, elsewhere in Svan, is attributable to paradigmatic levelling under the influence of the weak 1st conjugation. The syntax of activity verbs in Series II, likewise, has been 'borrowed' from the 1st conjugation, presumably because most of these verbs, although intransitive, have agents as subjects (in Harris's terms, there has been a shift from ergative to ACTIVE case marking).

1.3. Preverbed pseudo-aorists for Svan atelic verbs. I find the Kaldani/Harris hypothesis convincing in most respects. Where I differ from them is in the classification of forms such as *äd=šdir-al* in Prehistoric Svan. It is my belief that these were not originally aorists—or Series II screeves of any sort—but rather *Series I forms that were later reinterpreted as aorists*. The facts to support my view come from the Lower Bal subdialects. The imperfect stems of 3rd conjugation verbs in -*äl/el-* are marked by ablaut of the suffixal vowel to /*a*/, which Kaldani [1968] attributes to a now-lost imperfect-stem formant *-*a* [examples from Topuria 1967:90, 91 n.1]:

{14} Imperfect stems of 3rd conjugation verbs (Lower Bal)

| | | |
|----------------------|---------------------------------|--|
| S_{3sg} present: | i-šdr-äl 's/he plays' | i-burg-el 's/he wrestles' |
| S_{1sg} imperfect: | xw-i-šdr-al | xw-i-burg-al |
| S_{2sg} imperfect: | x-i-šdr-al | x-i-burg-al |
| S_{3sg} imperfect: | i-šdr-al \Leftarrow i-šdir-al | i-burg-al < *i-burg-el-a < *i-šdir-äl-a |

Now compare the above forms with the athematic aorists of the same verbs in the Laxamulan subdialect (the aorist of *išdräl* is in {13}, that for *iburgel* is given below):

{15} Athematic 'aorist' of 3rd conjugation verb (Lower Bal [Laxamulan])

| | |
|-----------|---|
| S_{1sg} | ot=burg-al \Leftarrow ad=xw-i-burg-al 's/he wrestled' |
| S_{2sg} | ät=burg-al \Leftarrow ad=x-i-burg-al |
| S_{3sg} | äd=burg-al \Leftarrow ad=i-burg-al |

The correlation is the same for all other activity verbs in *-äl/el-* which I tested with my informants. Some examples are given below, along with the Georgian glosses provided by the informants. In each case, the Laxamulan forms derived by the addition of a preverb to the imperfect were glossed by Georgian aorists:

{16} Imperfects and pseudo-aorists of 3rd conjugation verbs (Laxamulan)

| Imperfect | Pseudo-aorist | Georgian gloss of pseudo-aorist [S_{3sg} forms] | |
|----------------|----------------|--|----------------------------|
| i-k'il-al | äd=k'il-al | ik'ivla | 'yelled, shrieked' |
| i-q'ul-al | äd=q'ul-al | ibɣavla | 'mooed, lowed' |
| i-bk'ər-al | äd=bek'ur-al | ibluk'una | 'stammered' |
| i-k'ərək'ac-al | äd=k'ərək'c-al | ik'ak'ana | 'cackled' |
| i-k'wc-al | äd=k'wec-al | ik'vnesa | 'sighed' |
| i-c'kun-al | äd=c'kun-al | ic'veta | 'dripped _{intr} ' |

In the Laxamulan subdialect of Lower Bal, then, a sizeable number of atelic verbs—statives and activity verbs alike—have screeves with aorist-like semantics formed by adding a preverb (usually *ad-*) to the imperfect. The preverbed imperfects of activity verbs denote, according to my informants, actions which began and ended in the past, and which are represented as single events in the structuring of narrative (i.e., this screeve cannot be used when another event is described as occurring within the time frame of the activity in question). For some verbs, the preverbed imperfect conveys the additional sense of a *single occurrence* (semelfactivity), e.g., *eji äd=k'il-al* (\Leftarrow *ad=i-k'il-al*) [s/he-NOM yell:Pvb+Impf] 's/he yelled (once), uttered a shriek' (cf. imperfect *i-k'il-al*). Note also the following Lower Bal (Becho subdialect) forms, cited by Topuria [1967: 159, 166], who translates the Svan preverbed imperfect with a Georgian 2nd conjugation aorist (with punctiliar

semantics):

| | |
|---|---|
| {17} Imperfect | Pseudo-aorist |
| sip'-d-a | an=sip'-d-a |
| 'was turning' (Geo. <i>t'rialebda</i>) | 'turned (hither)' (Geo. <i>mot'rialda</i>) |
| Present Conjunctive | Pseudo-optative |
| sip'-d-e-s | an=sip'-d-e-s |

A similar form, based on the same root but with a different preverb, occurs in an Upper Bal poem [Šanidze et al 1939:150₈; recorded in the village Mulakhi in 1927]. Here as well the editors gloss the Svan form with a Georgian 2nd conjugation aorist:

- {18} čäž-ild mærtente lä=j-sip'-d-a.
 horse-DIM:NOM back-to turn:S_{3sg}:O₂:Pvb+Impf('AOR')
 'The horse turned around toward you.'
 Geo. *cxeni uk' u šegit' rialda*.

(These latter two examples indicate that preverbed imperfects of 3rd conjugation verbs with aorist-like semantics were not limited to the Laxamulan subdialect of Lower Bal.)

This phenomenon—the formation of a verb with semantics characteristic of Series II (i.e., punctiliar/semelfactive aspect) by the addition of a preverb to a Series I stem (imperfect or present conjunctive) appears to be limited to atelic verbs (3rd and 4th conjugations). On the other hand, the addition of a preverb to the imperfect or present conjunctive of a 1st or 2nd conjugation verb gives the (perfective) conditional and future conjunctive screeves, respectively, in both Georgian and Svan. The conditional is primarily used in (i.) hypothetical or counterfactual constructions ('had you called me sooner, I would have come [*conditional*] on time'); (ii.) as a future-in-the-past ('the day before yesterday I heard that he would buy a car [*conditional*] yesterday') [Vamling 1989:70]; and, (iii.) in descriptions of habitual past action ('as soon as the sun would rise, Peter would get up, dress, wash, eat breakfast, . . . ' [*all conditionals*]) [see Šanidze 1953:215; Mač'avariani 1974:120; Metreveli 1988]. Here is an example of a Svan perfective conditional, used to indicate the future-in-the-past in indirect speech:

{19} [Upper Bal; Šanidze et al 1939:38₅₁₋₅₂]

aš dam je:ka ugvarw čegmärs, gela kurda:n *ču=mi-c'wer-da!*

so not tell-S_{1sg}:O₂:AOR ignoble Chegem-PL-DATGela K.-NOM
 avenge-S_{3sg}:O_{1sg}:Pvb+Impf(PRFV.CONDL)

'Didn't I tell you low-life Chegemians that Gela Kurdan *would avenge me!*'

Geo. *xom gitxarit ugvaro čegemlebs gela kurdani ai yebda čems sisxls!*

The semantic range of these preverbed imperfects of Georgian and Svan telic verbs resembles that of the English and French constructions consisting of, formally speaking, the *past tense of the future* (for example, *would* + infinitive in English [*would* being the past-tense form of the future auxiliary *will*]; the French *conditionnel* [which bears the same morphological relation to the future as the imperfect to the present]). These facts can be represented with a sort of semantic algebra, as in {20}:

{20} **Bracketing of morphological components for Svan telic verbs.**

[preverb + present stem] + imperfect = future + imperfect = conditional,
 future-in-the-past, etc.

E.g., [*ču* + *mi-c'wer*] + *da* = *ču=mi-c'wer-da* 'he would avenge me'

By contrast, the bracketing associated with the pseudo-aorists of atelic verbs is as follows:⁵

⁵I would not exclude the possibility that the preverbed imperfects of at least some atelic verbs, even in Laxamulan Lower Bal, could be interpreted according to the bracketing pattern in {20} as well as {21}. In one of the Laxamulan texts in Davitiani et al. [1957:297₃₅₋₃₆], what would appear to be the preverbed imperfect of a 4th conjugation verb is used in the description of a hypothetical situation, a usage one would associate with the conditional:

| | | |
|-------------|---|---|
| šomä | mič | <i>es=c'ax-i-w,</i> |
| when | him/her:DAT | need:O ₃ :S _{3sg} :Pvb+Impf(PRFV.CONDL) |
| gam-s | xät'xa. | |
| in.kind-DAT | return:S _{3sg} :O ₃ :Impf | |

'If that person should ever be in need, (the family) would pay him back in

{21} **Bracketing of morphological components for Svan atelic verbs.**
 preverb + [present stem + imperfect] = perfective + past imperfect =
 perfective past (pseudo-aurist)

E.g.: $la + [xa-lat' + \text{ən-da}] = la=x-lat'-\text{ən-d-a}$ 'sb fell in love with sb'
 $ad + [i-\text{šdir-äl} + a] = \text{äd}=\text{šdir-al}$ 'sb played'

I will not dwell on these alternative bracketings any further in the present paper, save to point out that Svan morphological algebra does not abide by the associative law. I expect to return to this and other issues associated with the evolution of the Svan aspectual system in future work.

1.4. Aorists and pseudo-aorists of Svan atelic verbs. My hypothesis for the origin of the aorist- and optative-like screeves of Svan atelic (3rd and 4th conjugation) verbs is the following:

1. At an earlier stage of Prehistoric Svan, only telic (1st or 2nd conjugation) verbs had opposed Series I and Series II stems. The punctiliar aspect associated with Series II was of a sort that placed the narrative focus upon a salient *change of state*, and as such was incompatible with atelic verbs, the semantics of which represent an event or state as essentially homogeneous, without a significant transition point.

2. Preverbs in Prehistoric Svan came to be linked with *perfectivity* (the accomplishment or achievement of an event). Because of their semantics they became an obligatory component of the perfective screeves of Series I and both Series II screeves for telic verbs (except for some relics from the earlier aspectual system, in which preverbs were not used to code aspect).

3. Although their core significance is atelic, some events associated with 3rd and 4th conjugation verbs can be represented as punctiliar (e.g., single, brief occurrences: 'he uttered a cry') and/or telic (e.g.,

kind.'

The problem needs further study. One notes that the suffix *-w*, a common imperfect stem formant with most classes of verbs in Laxamulan [Topuria 1967:91 note 1], is replaced by *-ən* in most stative (4th conjugation) verbs. The Lower Bal pseudo-aorist given by Topuria [op cit, p 157] for this verb is $es=c'ex-\text{ən}$ (Gco *dasč'ir-da*) 'sb needed sthg'; hence the perfective conditional and pseudo-aorist are formally distinct, even though both consist of *preverb + present stem + imperfect suffix*.

the beginning of an event or state: 'he fell in love with her'). To convey such representations, Prehistoric Svan speakers added preverbs to certain Series I screeves of atelic verbs. Hence the preverbed imperfects and conjunctives given in {10} and {15}–{18} above.⁶

4. The descendents of the preverbed Series I screeves of stative verbs seem to have changed little in the modern Svan dialects: they have recognizable imperfect and present-conjunctive stems (note the presence of the suffixes *-ən-*, *-d-a*, *-de:d-* in {10}), which are never used in the Series II stems of telic verbs, but which do appear in the Series I stems of these verbs). As for the preverbed imperfects and conjunctives of activity verbs, these still retain their original morphological and syntactic characteristics in Laxamulan. Elsewhere they have taken on, in varying degrees, the formal properties of 1st or 2nd conjugation (true) aorists and optatives. In the table below are representative verbs from four inflectional patterns: 1st conjugation with weak aorist (*ama:re* 'sb prepares sthg'); 2nd conjugation with suffixal aorist (*imä:ri* 'sb/sthg gets prepared'); activity verbs in *-äl/əl-* which assign ERG case in Series II (*išdräl* 'sb plays'); activity verbs in *-äl/əl-* which assign NOM case in Series II (*idiräl* 'sb dines').

{22} Person-marking patterns for aorists in Upper Svan.

| Upper Bal | 3rd cj [+ERG] | 1st cj [weak] | 3rd cj [+NOM] | 2nd conj. |
|--------------------|---------------|---------------|---------------|------------|
| S _{2sg} : | ät=šdir-ä:l | ax-mä:r | ät=di:r-än | äx-mär-än |
| S _{3sg} : | äd=šdir-a:l-e | an-mär-e | äd=di:r-ä:n | än-mär-ä:n |
| Lower Bal | | | | |
| Becho | 3rd cj [+ERG] | 1st cj [weak] | 3rd cj [+NOM] | 2nd conj. |
| S _{2sg} : | ät=šdir-al | ax-mär | ät=dir-en | äx-mar-en |
| S _{3sg} : | äd=šdir-al-e | an-mar-e | äd=dir-an | än-mar-an |

⁶According to the Svan speakers I consulted, many root 3rd conjugation verbs (i.e., those with stems lacking the frequentative / durative suffixes *-ä:l-* or *-ie:l-*) form neither true aorist nor pseudo-aorist screeves. This is especially true in Lower Bal (with its more conservative aspect-marking morphology), and also in Upper Bal. Chato Gudjedjani provided over a dozen Upper Bal activity verbs the imperfects of which are not opposed to an aorist-like screeve, e.g. *ve:l-da* 'bleated,' *txi:l-da* 'snored,' *k'ə:r-a* 'shone,' *swib-da* 'danced.' Note also the lack of all preverbed screeves for some of the Lower Bal activity verbs listed in Gagua [1976:209-215].

| Ecer | <u>3rd cj [+ERG]</u> | <u>1st cj [weak]</u> | <u>3rd cj [+NOM]</u> | <u>2nd conj.</u> |
|--------------------|----------------------|----------------------|----------------------|------------------|
| S _{2sg} : | ät=šdir-äl | ax-mär | ät=dir-en | äx-mar-en |
| S _{3sg} : | äd=šdir-al-e | an-mar-e | äd=dir-än | än-mar-än |

| Laxamul | <u>3rd cj [+ERG]</u> | <u>1st cj [weak]</u> | <u>3rd cj [+NOM]</u> | <u>2nd conj.</u> |
|--------------------|----------------------|----------------------|----------------------|------------------|
| S _{2sg} : | ät=šdir-al | ax-mär | ät=dir-en | äx-mar-en |
| S _{3sg} : | äd=šdir-al-e | an-mar-e | äd=dir-an | än-mar-an |

Laxamul athematic aorist (= preverb + imperfect)

| | |
|--------------------|-------------------|
| S _{2sg} : | ät=šdir-al |
| S _{3sg} : | äd=šdir-al [+NOM] |

In Upper Bal, and in the Ecer subdialect of Lower Bal, 3rd conjugation verbs have gone completely over to one or the other of the patterns characteristic of telic verbs. In Becho and Laxamulan the S_{1/2sg} stem of the ERG-assigning 3rd conjugation lacks the umlaut which occurs in the 1st conjugation weak aorist.

5. These facts, in conjunction with the evidence presented by Kaldani [1978] and Harris [1985] which indicates the greater antiquity of the Laxamulan athematic aorist, lead me to reconstruct the following scenario for the development of 3rd conjugation aorists in Svan:

Scenario for history of 3rd conjugation verbs in Svan:

Stage I. 3rd conjugation verbs only form Series I screeves; they lack Series II screeves entirely. Their imperfect stems are formed by the addition of *-a in all persons and numbers (Kaldani [1968]; Mač'avariani [1980]). This suffix was later lost, after it lowered the /e/ of *-el* activity verbs to /a/ (cf. present *iburg-el* 's/he wrestles,' imperfect *iburg-al* in {14}). The preverbed imperfects and present-conjunctives assign NOM case to their subjects (*eji ad=i-burg-el* 's/he:NOM wrestled'), as is normal for *Series I* screeves.

Stage II. Accompanying a shift in the aspectual system of Prehistoric Svan, Series II ceases to be semantically incompatible with atelic verbs. The activity verbs in *-äl/el* take on Series II stems according to the patterns employed by telic verbs. Certain less agentive 3rd conjugation verbs (e.g., *i-rxun-äl* 'it thunders'; but note also the quite agentive *i-dir-äl* 'sb dines' in {22}) take on the inflectional characteristics of 2nd conjugation verbs in *-ä:n* (in the aorist series only). The rest form new screeves based on the 1st conjugation (along

with ERG case assignment), but retain their old 3rd conjugation pseudo-aorists (preverbed imperfects) as well.

| | <u>Old pseudo-aorist</u> | <u>1st cj pattern</u> | <u>2nd cj pattern</u> |
|----------------------|--------------------------|-----------------------|------------------------------|
| $S_{1/2sg}$: | -a:l | -ä:l | -än |
| $S_{3sg/all\ pls}$: | -a:l | -a:l-e | -ä:n |
| | [+nom] | [+erg] | [+nom; Cp. present in -ä:l] |

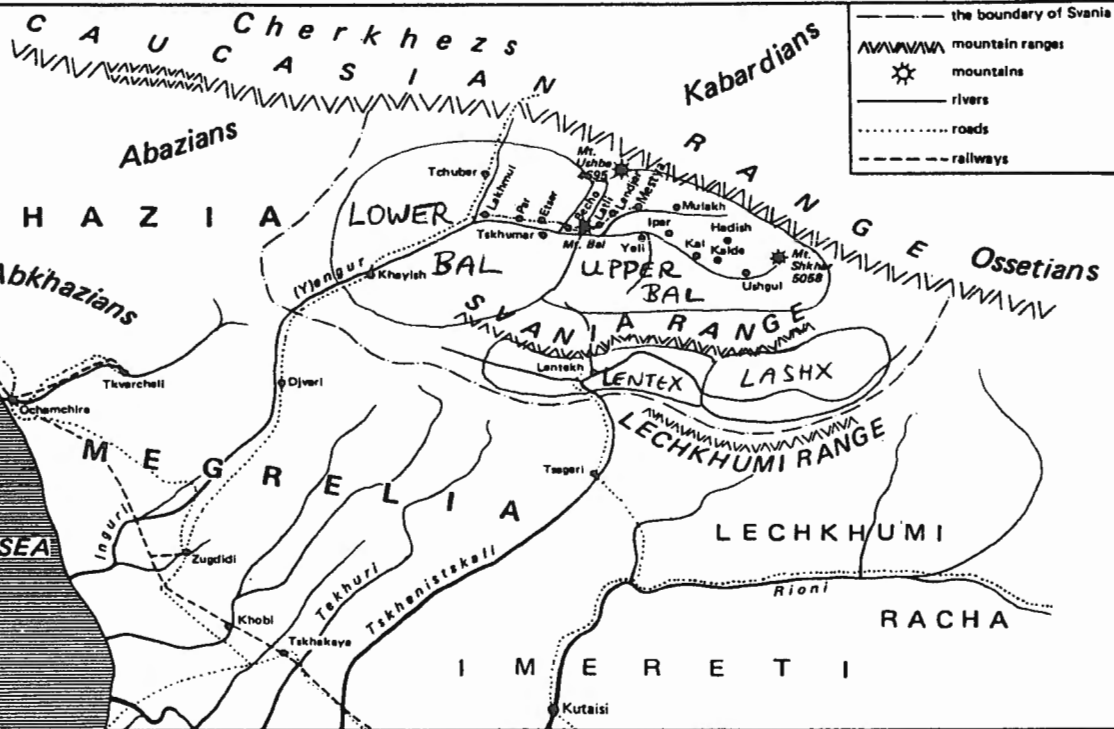
Stage III.

[**Upper Bal**] The old pseudo-aorist pattern is completely lost, replaced by a pattern (and case assignment) modeled after the 1st conjugation.

[**Lower Bal**] Traces of the old pseudo-aorist are retained to varying degrees, especially in the $S_{1/2sg}$ stem. In Laxamulan the pseudo-aorist and 1st-conjugation-based true aorist forms continue to survive side-by-side.

| | <u>Becho</u> | <u>Ecer</u> | <u>Laxamulan</u> | |
|----------------------|--|--|------------------|--|
| $S_{1/2sg}$: | -al | -äl/-al | [-al] | -al |
| $S_{3sg/all\ pls}$: | -al-e | -al-e | -al | -al-e |
| | | | [+nom] | [+erg] |
| | (1/2sg from ps-aor, 3sg/pls from 1cj) | (mostly 1cj, traces of ps-aor in 1/2sg) | (ps-aor) | (1/2sg from ps-aor, 3sg/pls from 1cj) |

Acknowledgments. The author gratefully acknowledges the assistance rendered by his father-in-law Ambak'o Č'k'adua (Toponymic Research Group, Tbilisi State University). Ambak'o provided useful criticism throughout the preparation of this paper, as well as much of the Laxamulan data cited. I am also thankful for the help in checking the Laxamulan facts provided by Ambak'o's brothers Arčil and Otar Č'k'adua (all born in the village Žabe [Upper] Ved, near Xaiši), Ek'at'erine K'axiani (born in Laxamula), and Givi Adzas-dze Č'k'adua (Axal Xaiši). Thanks also to Howard Aronson (University of Chicago) and Chato Gudjedjiani (Svanetian Ethnographic Museum, Mestia) for their comments on the ideas presented here. May Jgəräg forfend that any of them be held responsible for the errors contained in this paper.



The Svan dialects [map reproduced from Ch. Gudjedjiani & M. Palmitis *Upper Svan: Grammar and texts*. Vilnius: Mokslas, 1986; p. 15.]

REFERENCES.

- Č'umburidze, Zurab. 1986. *Mq'opadi kartvelur enebši* (The future tense in the Kartvelian languages). Tbilisi: Tbilisi State University Press.
- Davitiani, Aleks. 1973. *Svanuri andazebi* (Svan proverbs). Tbilisi: Mecniereba.
- Davitiani, A., V. Topuria and M. Kaldani, eds. 1957. *Svanuri p'rozauli t'ekst'ebi II: balskvemouri k'ilo* (Svan prose texts, II: Lower Bal dialect). Tbilisi: Mecniereba.
- Gagua, K'lara. 1976. *Dro-nak'li zmnebi svanurši*. (Defective verbs in Svan). Tbilisi: Mecniereba.
- Gudjedjiani, Chato & Mykolas L. Palmaitis. 1986. *Upper Svan: Grammar and texts*. Vilnius: Mokslas.
- Harris, Alice C. 1985. *Diachronic syntax: the Kartvelian case (Syntax and syntax 18)*. New York: Academic Press.
- Holisky, Dee Ann. 1981a. *Aspect and Georgian medial verbs*. Delmar, New York: Caravan Press.
- _____. 1981b. "Aspect theory and Georgian aspect" in *Tense and aspect (Syntax and semantics 14)*, ed. by P. Tedeschi and A. Zaenen, pp. 127-144. New York: Academic.
- Kaldani, Maksime. 1968. "E/a xmovantmonacvleobis zogi sak'itxisatvis svanur zmnebsi" (Some questions concerning the vowel alternation e/a in Svan verbs), *Iberiul-k'avk'asiuri enatmecniereba* 16:132-143
- _____. 1978. "Aorist'is c'armoeba svanurši" (Formation of the aorist in Svan), *Iberiul-k'avk'asiuri enatmecniereba* 20: 150-161.
- Mač'avariani, Givi. 1974. "Aspekt'is k'at'egoria kartvelur enebši" (The category of aspect in the Kartvelian languages), *Kartvelur enata st'rukt'uris sak'itxebi* 4:118-141.
- _____. 1980. "Namq'o usruli svanurši da misi adgili kartvelur enata uyvlilebis sist'emaši" (The past imperfect in Svan and its place in the conjugational system of the Kartvelian languages), *Iberiul-k'avk'asiuri enatmecniereba* 22: 207-218.
- Met'reveli, Teimuraz. 1988. "Asp'ekt'is šesc'avlis ist'oriidan kartulši" (From the history of the study of aspect in Georgian), *Dzveli kartuli enis k'atedris šromebi* 27:150-161.
- Q'ipšidze, Ioseb. 1914. *Grammatika mingrelskago (iverskago) jazyka s xrestomateju i slovarem. Materialy po jafetičeskomu jazykoznaniju* 8.
- Saxok'ia, Tedo. 1988. "Megruli andazebi da gamotkmebi" (Mingrelian proverbs and sayings), *Iberiul-k'avk'asiuri enatmecniereba* 28: 240-271.
- Schmidt, Karl Horst. 1984. "On aspect and tense in Old Georgian," *Folia Slavica* 7:290-302.
- Šanidze, Ak'ak'i. 1953. *Kartuli gramat'ik'is sapudzvlebi, I: morpologia* (The fundamentals of Georgian grammar I: morphology). Tbilisi: Tbilisi State University Press.
- Šanidze, A., V. Topuria and M. Gujejiani, eds. 1939. *Svanuri p'oezia I* (Svan

poetry, vol. I). Tbilisi: Mecniereba.

Topuria, Varlam. 1967. *Svanuri ena, I: zmna* (The Svan language, I: Verb). Tbilisi: Mecniereba.

Vamling, Karina. 1989. *Complementation in Georgian*. (*Travaux de l'Institut de Linguistique de Lund* 23). Lund University Press.

CODE SWITCHING AND THE ESTONIANS OF NARVA

Anna Verschik
Tallinn, Estonia

The town of Narva lies on the historical border between Estonia and Russia. It has been a multinational town throughout the centuries. During the interwar period of independence (1918-1940) the number of Estonians was never below 60%, the number of Russians was constantly 20-30 %, the rest of the town's population consisted of Germans, Finns, Jews, and others (Narva statistika aastramat 1932:4). The outcome of World War II and Soviet policy turned out to be fatal for the town, destroying its historical and cultural status and heritage. Most of the former citizens fled or were deported after the war. Today only 4-5% of the inhabitants are Estonians, while the rest are mainly Russian imports, i.e., speakers of Russian. The town has become culturally isolated from the rest of Estonia. The Estonians who still live there have to use Russian constantly. Russian is now the de facto official language in Narva, while Estonian is used mainly at home. It is quite difficult to receive education in Estonian: the town has only one school with Estonian the language of instruction, but even there several subjects are taught in Russian because of the lack of Estonian-speaking teachers.

When dealing with bilingualism two questions often emerge: is it possible to know both languages perfectly, "equally," and is a bilingual capable of keeping two languages apart? Society quite often differentiates the functions of two languages, i.e., each language is used in certain spheres or circumstances. In the town of Narva, Russian is the language of the official sphere, Estonian is used only at home and in the company of fellow Estonians; however, such a restricted usage of two languages is not fixed legally. Fishman (1972:45-48) assumes that from a sociolinguistical point of view a society does not need two languages in the same function.

To define a bilingual as a person who knows two languages "equally" well does not seem quite correct because this definition does not include, for example, immigrants, who know the local language poorly, yet are capable of elementary communication. If we claim as bilinguals only such persons who know both languages as well as a monolingual native speaker, the majority of people would belong to no group: they

are not monolingual, but according to the definition, are not bilingual either. Grosjean (1982:232) believes that this paradox confirms the need for a "more realistic definition" (1982:232): a bilingual is a person who regularly uses two languages in his everyday life (for more about various definitions of bilingualism see Skutnabb-Kangas 1983:2-18, 83 ff.).

Such a definition, based on the criterion of usage, allows a number of possibilities:

- (1.) people know both languages (and use them) from childhood, because their parents are of different ethnic origin and have spoken to them in their mother tongue;
- (2.) a person belongs to a minority group, using the childhood language at home and the language of the majority outside the home;
- (3.) a person is an immigrant and knows the language of the majority poorly;
- (4.) a person is a native of a colony and has to use two different languages—one with compatriots and another with the authorities.

The aforementioned implies that bilingualism cannot mean "equilingualism" because one uses two languages in different spheres and circumstances, especially in the case of diglossia (i.e., highly restricted differentiation of language usage).

It is, however, impossible to answer "yes" or "no" to the question, whether a bilingual is capable of keeping two languages apart. It happens that while communicating with another bilingual one switches from one language to the other quite frequently. In certain situations such a switching is a norm, not a violation of it. The alternative use of two languages in the same utterance or conversation is called code switching (Grosjean 1982:145). Code switching can be both conscious and unconscious. In the first case we deal with conscious language choice. A monolingual also chooses between styles, words, etc. within the same language; a bilingual chooses between two languages. In the case of diglossia the language choice is not free (Ferguson 1959:324-340, Grosjean 1982:129).

It turns out that the choice has two stages: first is the choice of language, second is the choice of code-switching or not code-switching; this latter choice is free (Grosjean 1982:129).

Let us list some of the reasons for code-switching: (1.) to fill in a "gap"; (2.) trigger-words (words which facilitate switching); (3.) quotation; (4.) to amplify of emphasis; (5.) to mark group or identity; (6.) to exclude someone from the conversation, etc. Gumperz (as summarized in Viikberg 1989:203) enumerates among the reasons for code-switching the following: at a certain moment code-switching contains important information, because choice of language is as significant as dialectal, stylistic, or any other choice. Even though no special research has been done on code switching of Estonians of Narva, we can describe some examples which come from the author's personal observation.¹ Sometimes it is quite difficult to decide whether we have a case of code switching or borrowing (if an item of the other language does not require any integration). The sentence *ma olen keroi* 'I am a hero' was produced by an informant who speaks Russian fluently, yet with an Estonian accent ([g'] > [k] at the beginning of the word). As the substantive *keroi* (literary Estonian *kangelane*) does not require any case marker in this position, it is impossible to decide if it is a nonce borrowing or code switching. If we consider the sentence to contain code switching, we shall mark the point of switching with //: *ma olen // keroi*. The most frequent reasons for code-switching among the Estonians of Narva are: (1.) mentioning certain topics and themes in the conversation (especially official matters, which are connected with use of Russian); (2.) to emphasize something; (3.) trigger words and sentences.

In the case of diglossia each language has its certain restricted function and field of usage. While discussing topics which deal with work, official documents, law, etc. (i.e., the sphere of Russian), switching to Russian is more than probable: *sinna Sillamäe linna ilma propuskita ei saa, seal on pograntsnik, // пограничник не пропустит* 'one can't go to the town of Sillamäe without written permission, there is a border guard, / the border guard would not let [one] through'; *Juta on*

¹The linguistic data come from the author's personal observation in 1987-1990. The author has spent every summer in Narva since 1968 and is acquainted with most of the informants from her childhood.

tööl // в пожарной команде // Narvas 'Juta works / at the fire-station / in Narva'.

Things said in a foreign language seem to be more powerful, especially when speaking to children, even if a child does not know any Russian: *räägi // tisse* 'speak / quietly'; *ära mine sinna, // не ходи туда* 'don't go there, / don't go there..

Sometimes one changes language in order to translate one's last words, that is, also emphasize them: *meil vett ei ole, ainult Mustjõe, aga see on nagu // капля в море, // nagu tilk meres* 'we have run out of water, we have only Mustjõe, but this is as / a drop of water in the sea, / as a drop of water in the sea'. Loans, international words, proper names can serve as trigger words: *on aga selline arganizatsija, // работать не хотят* 'such an institution (a literary Estonian organization), / they don't want to work'; *kui tuleb L'ena, // она скажет* 'when L'ena comes, / she will tell you.'

Types of code-switching (Grosjean 1982 : 320 ff.):

I. One-word or two-word switches:

1. Noun: *ma tahan olla // массажисткой* 'I want to be / a masseuse.'
2. Adjective: *nii on // солидной* 'that will be / more elegant.'
3. Adjective and nouns—no examples.
4. Verb: *mina // не умею // mängida* 'I / cannot / play.'

II. Longer switches:

1. Noun or verb phrase: *ära sinna mine, // не ходи туда* 'don't go there, / don't go there.'
2. Prepositional phrase: *elavad kuskil talus // под Нарвой* 'they live somewhere in a farmstead / near Narva.'
3. Clause: *ma ütlen talle, et // контора закрыта* 'I tell him that / the office is closed.'
4. Discourse markers, fillers : *ta tuli koju, // значит, // õhtul väga hilja* 'she came home, / that is, / very late in the evening.'
5. Idiomatic expressions: *ainult Mustjõe, aga see on nagu // капля в море* 'only Mustjõe, but it is / as a drop of water in the sea.'

Many linguists have started to recognize that code switching is linguistically constrained. In 1980 in a paper on English-Spanish bilingualism Poplack described the two following constraints (Poplack 1980:580-615, Viikberg 1989:204):

1. *Free-morpheme constraint*. A switch cannot take place between the stem of a word and its affix (bound morpheme). According to Poplack this constraint applies not only to affixes but also to idiomatic expressions. Viikberg, who has been studying the Estonians of Siberia (both the descendents of 19th century settlers and 20th century deportees), brings Estonian-Russian examples: there cannot exist such a form as **võlopetanu*, where *võ-* is a Russian prefix and *lopetanu* 'graduate' is an Estonian stem (cf. Russian выпускник 'graduate'); **utsiteljanna*, where *utsitelj* 'teacher' is a Russian stem and *-anna* is an Estonian marker of feminine gender.

2. *The equivalence constraint*. This requires that the word order immediately before and immediately after a switch point must be possible in both languages. In this case Viikberg could not provide examples, because he has not studied codeswitching in particular (Viikberg 1989:204). There are no other studies on Estonian-Russian bilingualism, so we can only suggest examples heard in Narva: *ma ütlen talle, et // контора закрыта, // mine koju*, cf. Estonian *ma ütlen talle, et / kontor on kinni, / mine koju* and Russian я говорю ему, что / контора закрыта, / иди домой 'I tell him, / that the office is closed, / go home'; *pane see ümber, nii on // солидней*; cf. Estonian *pane see ümber, nii on / soliidsem* and Russian обвяжи это вокруг, так будет / солидней 'bound it, it will be / more elegant'.

In addition to these two constraints, Timm and Gumperz (Grosjean 1982:327) suppose that code switching cannot take place between a subject pronoun and the verb. They also found that switching location between the verb and the infinitive complement is also very infrequent; the reason for this kind of restriction could be the tight semantic or prosodic bond between the two verbs. We claim that such code switching is not only possible, but also quite typical in Estonian-Russian bilingualism. One can register similar cases of code-switching also among bilinguals in other parts of Estonia, especially in Tallinn. Here are examples, where this restriction is frequently violated: *mina // ne умею // mängida* 'I / cannot / play'; *mina // не буду // laulda* ' / will

not / sing'; *mina* // *не хочу* // *tulla* 'I /don't want / to come'. All these sentences contain modal verbs and in each sentence the speaker tries to be most expressive.

There is no consensus among the scholars about the problem of the existence or nonexistence of a code-switch grammar. Some think that there are two monolingual grammars which can merge partly on the level of phrase structure; others find that there exists a code-switch grammar which is a combination of two lexicons and the grammatical categories of two languages and where certain restrictions are valid (Poplack 1980 : 580-615). We agree with Viikberg that one should first find out code-switching possibilities and restrictions in the case of Estonian-Russian bilingualism, and later try to explain the possibility of code-switching grammar.

REFERENCES

- Fergusson, C. A. 1959. Diglossia. *Word* 15:324-340.
- Fishman, J. A. 1972. The Sociology of Language. In: Giglioli, P. (ed.), *Language and Social Context*. Penguin Modern Sociology Readings. Middlesex, England, 45-48.
- Grosjean, F. 1982. *Life with Two Languages. An Introduction to Bilingualism*. Cambridge, Mass. Howard University Press.
- Narva statistika aastaraamat 1932. Narva.
- Poplack, S. 1980. Sometimes I begin a sentence in Spanish. *Linguistics* 18:580-615.
- Skutnabb-Kangas, T. 1983. *Bilingualism or Not: The Education of Minorities*. Multilingual Matters 7. Clevedon, Avon, England.
- Viikberg, J. 1989. Koodivahetus ja Siberi eestlased. *Lähtkohti*.—*Keel ja Kirjandus* 4:202-205.

ON ACCENT IN CHUVASH

Alexander Vovin
The University of Michigan

Among the many peculiarities that differentiate Chuvash from other Turkic languages, its accent system presents special interest. Various scholars have usually spoken about stress in Chuvash without identifying its nature (Andreev 1966:46), (Rezjukov 1959:47), (Jakovlev 1976, 160-161). One of the best native Chuvash linguists, V. G. Egorov, considered that Chuvash has a stress accent (Egorov 1954/1971:197). It seems to me that J. R. Krueger was the first scholar who suggested that Chuvash may have pitch accent (Krueger 1961:84).

According to my observations, made in different regions of the Chuvash republic in 1986-1988, Chuvash does not have a stress accent as in English or Russian, but a pitch accent as in Japanese or some Korean dialects.¹ Most of my materials come from the speakers of *anatri* ('lower') dialect, though I was able to elicit some phrases from the speakers of *viryal* ('upper') dialect as well.² In addition to these two basic dialects, there is also an intermediate 'middle lower' dialect which represents a kind of mixture of *anatri* and *viryal* dialects. I did not notice any significant difference in accentuation in these three dialects. I have not done any kind of experimental phonetic investigation of the Chuvash accent and thus the accent patterns below are established exclusively on the basis of my ear.

One interesting feature of the Chuvash accent is that while the other Turkic languages with pitch accent have only an oxytonic pattern (accent on the last syllable) with a few prototonic (accent on the first syllable) exceptions, Chuvash has three different patterns: oxytonic, prototonic, and mesotonic (accent on any medial syllable). However, patterns other than oxytonic are due to the presence in the word of the reduced vowels [ə] and [ɤ].³ These reduced vowels cannot take the

¹ By pitch accent I mean here merely a phonetic accent obtained by rise or fall of voice as opposed to a dynamic stress obtained by intensive articulation. From this point of view pitch accent characterizes not only standard or Kyoto Japanese, where the pitch accent is phonological, but also the "accentless" Japanese dialects, where the place of accent is fixed.

² My main informant was Piotr Vladimirovich Denisov, born in 1928 in the Batyrevo region in the south of the Chuvash republic, now a professor of history in the Cheboksary State University and the leading anthropologist in the republic.

³ Note that my notation of these vowels is the opposite of the traditional: [ə]

kernel of accent and so cause its shift. Chuvash accent follows the following rules:

(1) if there is no [ə] or [ǫ] in a word, it has the oxytonic pattern: *urá* 'foot,' *pukán* 'chair,' *śūxé* 'thin,' *sexét* 'watch,' *sulú* 'spoiling,' *kullén* 'every day,' *kuččenés* 'gift';

(2) if [ə] or [ǫ] are found in the last syllable, the accent shifts to the last syllable in the word which has vowels other than [ə] or [ǫ]: *álə* 'hand,' *śūrə* 'white,' *vísśǫ* 'three,' *xúləm* 'thick,' *ánsər* 'narrow,' *śéimǫs* 'green,' *śákə* 'this.'

(3) if the word contains only the vowels [ə] and [ǫ], the word has the prototonic pattern: *kǫkər* 'chest,' *ǫśə* 'warmth,' *śělēk* 'hat,' *kěřək* 'coat,' *śərəx* 'heat,' *xǫrlǫ* 'red,' *vǫrəm* 'long.'

On the basis of these rules, it should be possible to assume that different Chuvash accent patterns are secondary and their origin due to the reduction of vowels. However, there is an unusual correlation between the quantity of the vowel and the register of the Chuvash accent. I noticed three registers: high (H), medium (M), and low (L). The vowels [ə] and [ǫ] occur only in M and L registers, all other vowels only in H and M registers. Thus, we observe the following patterns:

1.1. *xir* H 'to rub'

1.2. *xǫr* M 'daughter'

2.1. *laša* MH 'horse'

2.2. *śūrə* HL 'white'

2.3. *kǫpe* LH 'shirt'

2.4. *kǫkər* ML 'chest'

3.1. *sapaka* MMH 'bunch'

3.2. *puslamǫś* MHL 'beginning'

3.3. *tǫtǫśən* MLL 'completely'

3.4. *kǫneke* LMH 'book'

3.5. *sivǫllǫ* HLL 'cold'

3.6. *pǫtrančək* LHL 'muddy'

3.7. *śǫtǫkla* LLH 'lace'

3.8. *ǫivǫxlan* MLH 'to approach'

It is obvious that when accented, [ə] and [ǫ] have middle pitch, and when unaccented, they have low pitch. Similarly, the other vowels are high-pitched when accented and middle-pitched when unaccented. Thus, denotes a back vowel, and [ǫ] a front vowel.

we can eliminate middle pitch as a superfluous phonological distinction and consider that the vowels [ə] and [ǝ] always bear low pitch, while all the other vowels bear high pitch. This provides us with the following accent classes:

| | | | |
|--------|---------|----------|----------|
| 1.1. H | 2.1. HH | 3.1. HHH | 3.5. HLL |
| 1.2. L | 2.2. HL | 3.2. HHL | 3.6. LHL |
| | 2.3. LH | 3.3. LLL | 3.7. LLH |
| | 2.4. LL | 3.4. LHH | 3.8. HLH |

Most 3-syllable classes are probably secondary, since the Chuvash nominal or verbal stem normally does not exceed two syllables (possible exceptions are seen in 3.1. and 3.7.); thus the unusual pitch pattern in 3.8. may be interpreted as HL-H.

I also establish one additional class, 1.2a (H < L-L), for one-syllable words, and two additional classes, 2.2a (HH < HL-L) and 2.4a (LH < LL-L), for two-syllable words. The words belonging to these classes look as if they belong to classes 1.1, 2.1, and 2.3 when used in isolation, but the genitive form demonstrates the original pattern.

Examples of the different accent classes:

1.1 H: *pin* 'thousand,' *jat* 'name,' *šin* 'person,' *kuś* 'eye,' *sūs* 'hair,' *puś* 'head,' *pit* 'face,' *ši* 'surface,' *kil* 'house,' *śurt* 'building,' *jen* 'side,' *jal* 'village,' *kaś* 'evening,' *ir* 'morning,' *xir* 'field.'

1.2 L: *pər* 'one,' *xər* 'girl,' *məj* 'neck,' *śəm* 'wool,' *mən* 'what,' *śər* 'land,' *xǝś* 'which,' *əs* 'mind,' *jər* 'trace,' *śəl* 'well,' *xəl* 'winter.'

1.2a H < L-L: *tu* H 'mountain' < *təv-ən* L-L (gen.), *śu* H 'butter' < *śəv-ən* L-L (gen.).

2.1 HH: *ura* 'foot,' *tumtir* 'clothes,' *xitre* 'beautiful,' *uram* 'street,' *atte* 'father,' *anne* 'mother,' *kukka* 'uncle (mother's brother),' *miše* 'how many/much,' *jultaś* 'friend,' *ača* 'child,' *muči* 'uncle (father's elder brother),' *śara* 'with bare feet,' *śemše* 'soft,' *laša* 'horse.'

2.2 HL: *ikkǝ* 'two,' *sakkər* 'eight,' *ivəl* 'son,' *arəm* 'wife,' *xuləm* 'thick,' *allə* 'fifty,' *kivǝ* 'old,' *atə* 'boots,' *tutər* 'shawl,' *alə* 'hand,' *pürnə* 'finger,' *xitə* 'hard,' *pisək* 'big,' *kantək* 'window.'

2.2a HH < HL-L: *širu* HH ‘letter’ < *širəv-ən* HL-L (gen.).

2.3 LH: *əšta* ‘where,’ *təvat* ‘four,’ *čəvaš* ‘Chuvash,’ *səxman* ‘type of upper garment (Russ. *kaftan*),’ *xəvel* ‘sun,’ *səran* ‘leather,’ *šəvar* ‘mouth,’ *čəlxe* ‘tongue,’ *tənče* ‘world,’ *əner* ‘yesterday,’ *səmax* ‘word,’ *pəri* ‘German wheat.’

2.4 LL: *təttəm* ‘dark,’ *vərə* ‘seed,’ *təxxər* ‘nine,’ *pətə* ‘porridge,’ *təsə* ‘color,’ *jəmək* ‘younger sister,’ *šəlləm* ‘younger brother,’ *šənə* ‘new,’ *pəčək* ‘small,’ *sələ* ‘oats,’ *jəvəs* ‘tree,’ *šəmər* ‘rain,’ *šənəx* ‘flour.’

2.4a LH < LL-L: *pəlü* LH ‘knowledge’ < *pələv=ən* LL-L (gen.), *kərü* LH ‘son-in-law’ < *kərv-ən* LL-L (gen.), *pəru* LH ‘calf’ < *pərv-ən* LL-L (gen.).

Since Chuvash pitch accent is connected with vowel quantity it is not phonological—there is no single pair of words which can be differentiated only by their accent patterns. However, I suggest that although purely phonetic, Chuvash pitch accent is very important for the reconstruction of Proto-Turkic vocalism.

Chuvash vocalism is famous for its chaotic correspondences with Common Turkic vocalism. Nevertheless, it is possible to establish regular correspondences for the Chuvash vowels [a], [u], [i], [ü], and [ɨ] with Common Turkic vowels, excluding loanwords from Tatar and several cases where the limited number of examples makes it difficult to show regular correspondences. However, Common Turkic correspondences for Chuvash [ə] and [ɤ] still remain in chaos—these vowels can correspond to almost any Common Turkic vowel.

I assume that these chaotic correspondences also may be explained if we project the Chuvash accent system onto the Common Turkic. In that case Chuvash [ə] and [ɤ] should be considered not as the reflexes of a specific Proto-Turkic vowel preserved in Common Turkic, but as the reflexes of a Proto-Turkic low pitch. Thus, Common Turkic lost the original accent patterns, while preserving a vowel system closer to the archetype than the present-day Chuvash vowel system. Meanwhile, the original low pitch in Proto-Chuvash suppressed vowel length,

producing reduced vowels, at the same time Chuvash still exhibits indirect evidence of the existence of pitch accent in Proto-Turkic.

Below I provide a new version of a reconstruction of Proto-Turkic vocalism which is based on the above assumption that Proto-Turkic had pitch accent. For the classification of Turkic languages I follow the point of view proposed by Z. Gombocz that Turkic may be divided into three main branches: Common Turkic including the majority of the Turkic languages, Yakut, and Chuvash (Gombocz 1912). That means that a reconstruction of Proto-Turkic vocalism should be based on the comparison of representatives from each branch. I usually use Turkish as a representative of Common Turkic, since the Oghuz languages preserved Common Turkic vocalism better than other subgroups of Common Turkic. If a word is not present in Turkish, I cite it from Turkmen or Old Osman. At the same time I use Old Turkic data with great care since this language has some innovative vowel shifts.

Let us look at vowel systems in our sample languages—Turkic, Chuvash and Yakut.

Turkish vocalism:

| | | | |
|---|---|---|---|
| ü | i | ı | u |
| | o | e | o |
| | | a | |

Yakut vocalism:

| | | | |
|---|---|---|---|
| ü | i | ı | u |
| | ö | e | o |
| | | a | |

Chuvash vocalism:

| | | | |
|---|---|----|----|
| ü | i | ı | u |
| | e | ə̌ | ə̍ |
| | | a | |

I have excluded from the comparisons Chuvash [e] as a kind of new phoneme, since it almost never occurs in the first syllable of words which have parallels in the other Turkic languages. I assume that the Chuvash vowels [ə̌] and [ə̍] correspond to Proto-Turkic low-pitched vowels while [a], [ı], [i], [u] and [ü] correspond to Proto-Turkic high-pitched vowels. Thus, Yakut vocalism is closer to the Common Turkic vocalism than to Chuvash vocalism, and, as far as I can judge, Yakut does not have pitch distinctions like Chuvash. In some cases, however,

it supports Chuvash reflexes of Proto-Turkic high-pitched vowels versus reflexes in Common Turkic as reflexes of the special Proto-Turkic high-pitched vowels which fell together in Common Turkic.

Correspondences for PT high-pitched vowels

| PT ⁴ | TUR | CHUV | YAK |
|-----------------|-------|-------|-----|
| *a | a | ǎ | ǎ |
| *ɑ | a | a | a/ǎ |
| *ɒ | a | u | a |
| *i | a/ǎ/i | ǎ/i | ǎ |
| *ə | ö | u | ö |
| *i | i/e | i | i |
| *ü | ü | ü/vi- | ü |
| *e | e | i | e |
| *ö | ö | ü/vi- | ö |
| *ε | e | a | e |
| *u | u | u | u |
| *u | u | ǎ | u |
| *o | o | u/vu- | o |

From this chart we can see that the reliability of a PT reconstruction based almost exclusively on the Common Turkic data is quite limited: Common Turkic does not differentiate PT *a, PT *ɒ, and PT *ɑ, and such a detailed reconstruction is also partly supported by Yakut data. Common Turkic did not preserve the distinction between PT *u and PT *u, PT *e and PT *ε, PT *ö and PT *ə.

Correspondence: TUR [a] : CHUV [ǎ] : YAK [ǎ] < PT *a.

PT *damar HH 'vein,' 'sinew' > TUR *damar*, CHUV *timar*, YAK *timir*.

PT *taj H 'foal' > TUR *tay*, CHUV *hijxa*, YAK *hij*.

PT *kata HL 'strong,' 'hard' > TUR *kati*, CHUV *xitə*, YAK *kitaanax*.

⁴ Abbreviations used are: CT—Common Turkic, CHUV—Chuvash, H—high pitch, J—Japanese, L—low pitch, MSFOu—Mémoires de la Société Finno-Ougrienne, OJ—Old Japanese, OT—Old Turkic, PJ—Proto-Japanese, PM—Proto-Mongolian, PMT—Proto-Manchu-Tungus, PT—Proto-Turkic, TUR—Turkish, TURK—Turkmen, YAK—Yakut

- PT **al*= H 'to take' > TUR *al*=, CHUV *-il=maš*=, YAK *-il*=.
- PT **jar*₁ H 'ravine' > TUR *yar* 'precipice,' CHUV *šir[an]*, YAK *sir* 'ravine.'
- PT **jat*= H 'to lie [down]' > TUR *yat*=, YAK *šit*= 'to lie [down],'
CHUV *ššar* 'pillow.'
- PT **ar₁ik* HH 'lean' > TUR *arık*, CHUV *-ırxan*, YAK *-ırigan*.
- PT **agr[a]*= H[H] 'to be sick' > TUR *ağrı*= 'to ache,' CHUV *-irat*=,
YAK *-iūarat* 'to get sick.'
- PT **ajit*= HH 'to ask' > OT *ajit*= 'to let/make ask,' CHUV *-ijt*=, YAK
-ijit= 'to ask.'
- PT **al₁tin* HL 'gold' > TUR *altın*, CHUV *-iltən* 'gold,' YAK *-iltahin*
'tin.'
- PT **aagır*₁ HHL 'heavy' > TUR *ağır*, CHUV *jivər*, YAK *-iar*.

Correspondence: TUR [a] : CHUV [a] : YAK [a/ä] < PT *a.

- PT **anne* HH 'mother' > TUR *analanne*, CHUV *anne* (no YAK cognates).
- PT **ana* HH 'hole' > CHUV *ana*, YAK *an* (no Common Turkic cognates).
- PT **atte* HH 'father' > TUR *ata*, CHUV *atte* (no YAK cognates).
- PT **aš* H 'food' > TUR *aş* 'food,' CHUV *aš* 'meat,' YAK *-ışik* 'food.'
- PT **aat* HH 'name' > TUR *ad*, CHUV *jat*, YAK *aat*

Correspondence TUR [a] : CHUV [u] : YAK [a] < PT *a.⁵

- PT **at* H 'horse' > TUR *at*, CHUV *ut*, YAK *at*.
- PT **ad₁v=k* HH 'leg,' 'foot' > TUR *ayak*, CHUV *ura*, YAK *atax*.
- PT **ad₂č* H 'head' > TUR *baş*, CHUV *puš*, YAK *bas*.
- PT **kav₁a* HH 'black.' TUR *kara*, CHUV *xura*, YAK *xara*.
- PT **ad₁[v]i=k* H[H]L 'fish' > TUR *balık*, CHUV *pulə*, YAK *bañiq*.
- PT **ad₂č* HH 'stone' > TUR *taş* (Old Osman *daş*), CHUV *čul*, YAK
taas.
- PT **jov₂* HH 'spring' > TUR *yaz*, CHUV *šur*, YAK *saas*.
- PT **jog[i]r(=)in* H[L]L 'back' > TUR *yağır* 'withers,' CHUV *šurəm*
'back,' YAK *šarın* 'shoulder.'
- PT **pl₁i* HL 'six' > TUR *altı*, CHUV *ultə/ultə/ult*, YAK *alta*.
- PT **kd₁iŋ* HL 'birch' > TUR *kayın*, CHUV *xurən*, YAK *xatŋ*.

⁵ I find the standard point of view that Chuvash [u] < CT [a] untenable, because [a] in CT may correspond to other Chuvash phonemes.

PT **ḍč*= H 'to open' > TUR *aç*=, CHUV *uś*=, YAK *as*=.

Correspondence TUR [a/ɨ/i] : CHUV [ɨ/i] : YAK [ɨ] < PT *ɨ

PT **igač* HL 'tree' > TUR *ağaç* (OT *igač*), CHUV *jivəs* (no YAK cognates).

PT **itV* HL 'dog' > TUR *it* (OT *it/it*), CHUV *jitəl/jit*, YAK *it*.

PT **kir*₁ H 'field,' 'steppe' > TUR *kır*, CHUV *xir* (no YAK cognates).

PT **šiš*= H 'to swell' > TUR *şiş*=, CHUV *šiš*= (no YAK cognates).

PT **bi[j]iV* HL 'louse' > TUR *bit*, CHUV *pjītə*, YAK *bit*.

Correspondence TUR [ö] : CHUV [u] : YAK [ö] < PT *ə.

PT **kəč*= H 'to cross,' 'to move' > TUR *göç*=, CHUV *kuś*=, YAK *kös*=.

PT **gər₂č* H 'eye' > TUR *göz*, CHUV *kuś* (no YAK cognates).

PT **gər₁*= H 'to see' > TUR *gör*=, CHUV *kur*= (no YAK cognates).

PT **tərV* H? 'relatives,' 'relationship' > TUV *törü*= 'to give birth,' CHUV *turtu* 'relatives,' 'relationship,' YAK *törüt* 'ancestor,' 'clan.'

Correspondence: TUR [i/e] : CHUV [i] : YAK [i] < PT *i.

PT **bir₂* H 'we' > TUR *biz*, CHUV *e=pir*, YAK *bisigi*.

PT **biil₁* HH 'waist' > Old Osman *bil* (TUR *bel*, TURK *biil*), CHUV *pilə̌k*, YAK *biil*.

PT **kir* H 'dirt' > TUR *kir*, YAK *kir* 'dirt,' CHUV *kirə̌k* 'dandruff,' 'dirt on the body or clothes.'

PT **il₂[i]t*= H[L?] 'to hear' > TUR *işit*=, CHUV *ilt*=, YAK *ihit*=.

PT **dir₂*= H 'to thread' > TUR *diz*=, CHUV *tir*=, YAK *tis*=.

PT **tik*= H 'to stab,' 'to stick' > TUR *dik*=, CHUV *čik*= 'to stab,' 'to stick,' YAK *tik*= 'to sting.'

PT **tir₁i* H[L?] 'skin' > TUR *deri*, CHUV *tir*, YAK *tiri*.

PT **ir₁i*= HL 'to melt' > TUR *eri*= CHUV *irə̌l*=, YAK *ir*=.

Correspondence TUR [ü] : CHUV [vi-/ü-] : YAK [ü] < PT *ü.

PT **tüg=me* H-H 'lace' > TUR *düğme* 'button,' CHUV *tüme* 'button,' 'lace,' YAK *tüm*= 'to knot.'

PT **tür₂V* HL 'straight,' 'honest' > TUR *düz*, CHUV *türə̌* 'straight,' YAK *tüs* 'serious.'

PT **jür₁e*= HH 'to go,' 'to move' > TUR *yürü*=, CHUV *šüüre*= (no

YAK cognates).

PT **üč*V HL 'three' > TUR *üç*, CHUV *viššă/višă/viš*, YAK *üs*.

PT **bür*₁V HH 'to wrap up' > TUR *bürü*=, CHUV *pürke*= (no YAK cognates).

Correspondence: TUR [e] : CHUV [i] : YAK [e] < PT **e*.

PT **gel*₁ H 'to come' > TUR *gel*=, CHUV *kil*=, YAK *kel*=.

PT **jedi* HL 'seven' > TUR *yedi*, CHUV *šičă/šičăčă*, YAK *sette*.

PT **je* H 'to eat' > TUR *ye*=, CHUV *ši*=, YAK *sie*=.

PT **bel*₂ H 'five' > TUR *beş*, CHUV *pilăk/pillăk*, YAK *bies*.

PT **bej*=*ni* HH 'brain' > TUR *beyin*, CHUV *mime*, YAK *mejii*.

PT **jel* H 'mane' > TUR *jele*, CHUV *šilxe*, YAK *siel*.

PT **em* H 'medicine' > TUR *em*, CHUV *im*, YAK *em*.

Correspondence TUR [ö] : CHUV [vi-/ü] : YAK [ö] < PT **ö*.

PT **köl*₁V HL 'lake' > TUR *göl*, CHUV *külă*, YAK *küöl*.

PT **öpke* HH 'lungs' > OT *öpke*, CHUV *üpke* (no YAK cognates).

PT **öl*= H 'to die' > TUR *öl*=, CHUV *vil*=, YAK *öl*=.

PT **ört* H 'fire (disaster)' > OT *ört*, CHUV *virt*, YAK *ört*.

PT **ölč*= H 'to measure' > TUR *ölç*=, CHUV *viş*= (no YAK cognates).

PT **böwre*=*k* HH 'kidney' > TUR *böbrek*, CHUV *püre*, YAK *büör*.

Correspondence: TUR [e] : CHUV [a] : YAK [e] < PT **ę*.

PT **ęl*₁i HL 'arm,' 'hand' > TUR *el*, CHUV *ală*, YAK *elii*.

PT **er*₁ H 'man,' 'male' > TUR *er*, CHUV *ar*, YAK *er*.

PT **geęž*[*ę*] 'night,' 'evening' > TUR *gece*, CHUV *kaś*, YAK *kiese*.

PT **bęer*₁= HH 'to give' > TUR *ver*= CHUV *par*=, YAK *bier*=.

PT **geč*= H 'to pass,' 'to cross' > TUR *geç*=, CHUV *kaś*=, YAK *kes*=.

PT **tep*= H 'to stamp one's foot' > TUR *tep*=, CHUV *tap*=, YAK *tep*=.

PT **deę* H 'same' > TURKM *deę*, CHUV *tan*, YAK *teę*.

Correspondence TUR [u] : CHUV [u] : YAK [u] < PT **u*.

PT **jurt* H 'camp' > TUR *jurt* 'country,' 'motherland,' CHUV *šurt* 'house,' YAK *surt* 'camp,' 'country.'

PT **ul*V= HH 'to howl' > TUR *ulu*=, CHUV *ula*=, YAK *uluj*=.

PT **juuw*= HL 'to wash' > TURKM *juv*= (OT *juu*=), CHUV *šu*=/*šăv*=,

YAK *suuj*=.

PT **kurt* H 'worm' > TUR *kurt*, CHUV *xurt* (no YAK cognates).

PT **kur*₁= H 'to build' > TUR *kur*= 'to build,' CHUV *xuralt(ə)* 'building' (no YAK cognates).

PT **kum* H 'wave' > OT *qum*, CHUV *xum* (no YAK cognates).

Correspondence TUR [u] : CHUV [ɨ] : YAK [u] < PT *u.

PT **šuw* H 'water' > TUR *su* (OT *sub*), CHUV *šiv*, YAK *uu*.

PT **tut*= H 'to hold' > TUR *tut*=, CHUV *hit*=, YAK *tut*=.

PT **ujqu* HL 'sleep' > TUR *uyku*, CHUV *-ijxə* 'sleep,' YAK *uuk* 'sleepy.'

Correspondence TUR [o] : CHUV [u] : YAK [o] < PT *o.

PT **oqV* HL 'arrow' > TUR *ok*, YAK *ox* 'arrow,' CHUV *uxə* 'bow.'

PT **og*= H 'to rub' > TUR *oğ*=, CHUV *ux*= (no YAK cognates).

PT **og* (**ög* ?) H 'thought' > OT *öγ*, CHUV *uj*, YAK *oj*.

PT **oj* H 'cavity,' 'valley' > TUR *oyuk* 'cavity,' CHUV *uj* 'field' (no YAK cognates).

PT **kol*₁ H 'arm' > TUR *kol*, CHUV *xul* (no YAK cognates).

PT **jol*₁ H 'way' > TUR *jol*, CHUV *śul*, YAK *suol*.

PT **oot* HH 'fire' > TUR *ot*, CHUV *vut*, YAK *uot*.

PT **bol*₁= H 'to be,' 'to become' > TURKM *bol*=, CHUV *pul*=, YAK *buol*=.

Correspondences for PT low-pitched vowels

| <u>PT</u> | <u>TUR</u> | <u>CHUV</u> | <u>YAK</u> |
|-----------|------------|-------------|------------|
| *a | a/e | ə | a |
| *ɒ | a | ə | a |
| *i | -i/i | ə | -i |
| *u | u | ə | u |
| *o | o | ə | o |
| *ə | ö | ə | ö |
| *i | i/e | ə | i |
| *ü | u | ə | ü |

Correspondence TUR [a/e] : CHUV [ə] : YAK [a] < PT *a.

PT **bagVr*₁ LH 'liver' > TUR *bağır*, CHUV *pəver*, YAK *biar*.

PT **kawVš*= LL > OT *qav-ış*= 'to scold each other,' 'to quarrel,'
CHUV *xəvəš*= 'to go back and forth,' 'to make a fuss' (no YAK cognates).

PT **jaŋi* LL 'new' > TUR *yeni*, CHUV *šənə*, YAK *saŋa*.

Correspondence TUR [a] : CHUV [ə] : YAK [a] < PT *a.

PT **kəčan* LH 'when' > TURKM *xačan*, CHUV *xəšan*, YAK *xahan*.

PT **təguk* LL 'chicken' > TUR *tavuk*, CHUV *čəxə* (no YAK cognates).

PT **bək*= L 'to look' > TUR *bak*=, CHUV *pəx*= (no YAK cognates).

PT **dəgər* LL 'stallion' > TUR *aygır*, CHUV *əjər*, YAK *atır*.

Correspondence TUR [ɛ/i] : CHUV [ə] : YAK [i] < PT *i.

PT **kil*₁ L 'horse hair' > TUR *kıl*, CHUV *xələx*, YAK *kil*.

PT **gıl*₂ L 'winter' > TUR *kış*, CHUV *xəl*, YAK *kihin*.

PT **il*₂ L 'work' > OT -š (TUR *ış*), CHUV *ış* (no YAK cognates).

PT **dil*₁ L 'tongue' > TUR *dil*, CHUV *čəlxə*, YAK *hil*.

Correspondence TUR [u] : CHUV [ə] : YAK [u] < PT *u

PT **tuCVr*₂ LH 'salt' > TUR *tuz*, CHUV *təvar*, YAK *tuus*.

PT **buur*₂ LL 'ice' > TUR *buz*, CHUV *pər*, YAK *muus*.

PT **bur*₂ LH 'calf' > TUR *buzağı*, CHUV *pəru* (no YAK cognates).

Correspondence TUR [o] : CHUV [ə] : YAK [o] < PT *o

PT **boj*=Vn L -? 'neck' > TUR *boyun*, CHUV *məj*, YAK *mooj*.

PT **kon*=Vk L-H 'guest' > TUR *konuk*, CHUV *xəna*, YAK *xonos*.

PT **kor*₁kV= LH 'to be afraid' > TUR *kork*=, CHUV *xəra*=, YAK *xork*=.

Correspondence TUR [ö] : CHUV [ə] : YAK [ö/o] < PT *ä.

PT **göwVk* LL 'blue' > TUR *gök*, CHUV *kəvak*, YAK *küöx*.

PT **gök* L 'root' > TUR *kök*, CHUV *kək* (no YAK cognates).

PT **tök*= L 'to pour' > TUR *dök*=, CHUV *təx*=, YAK *tox*=.

Correspondence TUR [i] : CHUV [ə] : YAK [i] < PT *i.

PT **ič*= L 'to drink' > TUR *iç*=, CHUV *ış*=, YAK *is*=.

PT **bil*₁= L 'to know' > TUR *bil*=, CHUV *pəl*=, YAK *bil*=.

PT **bil*V[w/g] LL 'whetstone' > TURKM *bilev*, CHUV *pəlev* (no YAK cognates).

Correspondence TUR [ü] : CHUV [ə] : YAK [ü] < PT *ü.

PT **jür*₂ L 'hundred' > TUR *yüz*, CHUV *śər*, YAK *süüs*.

PT **čjür*₁=Vk L-H 'heart' > TUR *yürek*, CHUV *čəre*, YAK *surex*.

PT **gü[ü]r*₂ L[L] 'fall' > TUR *güz*, CHUV *kər*, YAK *küüs*.

It is easy to notice that it is not possible to reconstruct the same number of low-pitched vowels in comparison with high-pitched vowels. This seems to be natural since in many cases where Common Turkic and Yakut reflexes of Proto-Turkic vowels fell together, only Chuvash reflexes provide the evidence for the reconstruction. But in the case of low-pitched vowels, Chuvash does not show any specific reflexes except for the (+back) or (-back) nature of the vowel with one exception: PT *ə is reflected in Chuvash as [ə] as if it were a back vowel. Thus, the full picture of the Proto-Turkic vocalism may be reconstructed only for high-pitched vowels.

Summarizing above correspondences for high-pitched vowels, we get the following system of the Proto-Turkic vocalism:

| | | | | |
|---|---|---|---|---|
| ü | i | ı | u | u |
| ö | e | ə | | o |
| | ε | | | |
| | a | ɑ | ɒ | |

The system above has one obvious deficiency - it has three low vowels [a], [ɑ], and [ɒ]. This system will look more natural and symmetrical if we raise vowel [ɒ], making it [ɔ]:

| | | | | |
|---|---|---|---|---|
| ü | i | ı | u | u |
| ö | e | ə | | o |
| | ε | | | ɔ |
| | a | ɑ | | |

The proposed reconstruction is not by any means the last word in Turkic comparative linguistics. There are still many unsolved problems, such as the coexistence of two high non-front unrounded vowels /ɨ/ and /u/, which seems to be unnatural from a phonological point of

view. I do not know how to resolve this problem at the present stage; possibly these two phonemes are really the same, especially taking into consideration the rare occurrence of PT *u. However, I believe that this reconstruction involving pitch distinctions in Proto-Turkic explains chaotic correspondences between Common Turkic vocalism, on the one hand, and Chuvash vocalism, on the other hand, considerably better than any previous attempts (Poppe 1960:91-117, Ščerbak 1970:76-77, Illič-Svityč 1971:171, Tenishev 1984:65-71, Serebrennikov and Gadžieva 1986:8-27).

I would also like to cite several parallels from other Altaic languages, where, it seems, the present reconstruction made these comparisons look better.

(1) G. Ramstedt rejected a comparison between CT *qara* 'black' and J *kuro* = 'black' on the basis of different vowels in the first syllable (Ramstedt 1924/1951:14). J *kuro* = is attested in OJ as *kurwo* = which suggests PJ **kuraCu* =. According to my reconstruction, the PT form should be **kɔra* or **kɔra*. The PJ vowel system is very poor (only *a, *i, *u, *ə) in comparison with the PT vowel system. Thus, if we assume that in pre-PJ there was a merging of PA *u and *o (and possibly *ɔ) into *u, then we get a perfect cognate set.

(2) PT **kata* 'strong,' 'hard' definitely has an [a] in the first syllable and here it is paralleled by PMT **kata* id., PM **kata* = id., PJ **kata* = 'hard,' also with a distinctive [a].

(3) A comparison of CT *taaš* 'stone' with MK *twol* 'stone,' PJ **[d]iso=Ci* and PMT **žolo* 'stone' is rejected by many opponents of the Altaic theory on the basis of different vowels in CT and the other languages in question. However, PT **dɔɔl₂* 'stone' agrees perfectly with the rounded vowels in MK and PMT. I suggest PA **djool₂o* or **d'ool₂o* 'stone.'

REFERENCES

- Andreev, I. A. 1966. "Čuvašskij jazyk." In *Jazyki narodov SSSR*, v. 2, pp. 43-65. Moskva.
- Egorov, V. G. 1954/1971. *Sovremennij čuvašskij literaturnyj jazyk v sravnitel'no-istoričeskom osveščanii*. Čeboksary.
- Gombocz, Z. 1912. *Die bulgarisch-türkischen Lehnwörter in der ungarischen Sprache*. MSFOu XXX.
- Jakovlev, P. Ja. 1976. "Udarenie v russkix i internacional'nyx slovaх,

- zaimstvovannyx v čuvaškij jazyk." In *Trudy naučno-issledovatel'skogo instituta pri Sovete Ministrov Čuvaškoi SSSR*. No. 66, pp. 159-175. Čeboksary.
- Illič-Svityč, V. M. 1971. *Opyt sravneniia nostratičeskix jazykov*. v. 1. Moskva.
- Krueger, J. R. 1961. *Chuvash Manual*. *Indiana University Publications, Uralic and Altaic Series*, vol. 7. Indiana University, Bloomington.
- Poppe, N. 1960. *Vergleichende Grammatik der Altaischen Sprachen. Teil 1. Vergleichende Lautlehre*. Otto Harrassowitz. Wiesbaden.
- Ramstedt, G. J. 1924/1951. "A Comparison of the Altaic Languages with Japanese". *Journal de la Société Finno-Ougrienne*, 55, pp. 7-24, Helsinki.
- Rezjukov, N. A. 1959. *Sopostavitel'naja grammatika russkogo i čuvaškogo jazykov*. Čeboksary.
- Ščerbak, A.M. 1970. *Sravnitel'naia fonetika tjurkskikh jazykov*. Leningrad.
- Serebrennikov, B. A. and Gadžieva, N. Z. 1986. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov*. Moskva.
- Tenišev, E. R. 1984. "Pratjurskij vokalizm. Reprezentacija pratjurskix glasnyx v sovremennyx jazykax i dialektax." In: *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Fonetika*. Moskva.